

*Citation for published version:*

Komjáthy, J & Lodewijks, B 2020, 'Explosion in weighted hyperbolic random graphs and geometric inhomogeneous random graphs', *Stochastic Processes and their Applications*, vol. 130, no. 3, pp. 1309-1367. <https://doi.org/10.1016/j.spa.2019.04.014>

*DOI:*

[10.1016/j.spa.2019.04.014](https://doi.org/10.1016/j.spa.2019.04.014)

*Publication date:*

2020

*Document Version*

Peer reviewed version

[Link to publication](#)

*Publisher Rights*

CC BY-NC-ND

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# EXPLOSION IN WEIGHTED HYPERBOLIC RANDOM GRAPHS AND GEOMETRIC INHOMOGENEOUS RANDOM GRAPHS

JÚLIA KOMJÁTHY AND BAS LODEWIJKS

**ABSTRACT.** In this paper we study weighted distances in scale-free spatial network models: hyperbolic random graphs, geometric inhomogeneous random graphs and scale-free percolation. In hyperbolic random graphs,  $n = \Theta(e^{R/2})$  vertices are sampled independently from the hyperbolic disk with radius  $R$  and two vertices are connected either when they are within hyperbolic distance  $R$ , or independently with a probability depending on the hyperbolic distance. In geometric inhomogeneous random graphs, and in scale-free percolation, each vertex is given an independent weight and location from an underlying measured metric space and  $\mathbb{Z}^d$ , respectively, and two vertices are connected independently with a probability that is a function of their distance and their weights. We assign independent and identically distributed (i.i.d.) weights to the edges of the obtained random graphs, and investigate the weighted distance (the length of the shortest weighted path) between two uniformly chosen vertices, called typical weighted distance. In scale-free percolation, we study the weighted distance from the origin of vertex-sequences with norm tending to infinity.

In particular, we study the case when the parameters are so that the degree distribution in the graph follows a power law with exponent  $\tau \in (2, 3)$  (infinite variance), and the edge-weight distribution is such that it produces an *explosive* age-dependent branching process with power-law offspring distribution, that is, the branching process produces infinitely many individuals in finite time. We show that in all three models, typical distances within the giant/infinite component *converge in distribution*, that is, no re-scaling is necessary. This solves an open question in [51].

The main tools of our proof are to develop elaborate couplings of the models to infinite versions, to follow the shortest paths to infinity and then to connect these paths by using weight-dependent percolation on the graphs, that is, we delete edges attached to vertices with higher weight with higher probability. We realise the percolation using the edge-weights: only very short edges connected to high weight vertices are allowed to stay, hence establishing arbitrarily short upper bounds for connections.

## 1. INTRODUCTION

Many complex systems in our world can be described in an abstract way by networks and are studied in various fields. Examples are social networks, the Internet, the World Wide Web (WWW), biological networks like ecological networks, protein-protein interaction networks and the brain, infrastructure networks and many more [59, 68]. We often do not comprehend these networks at all in terms of their topology, mostly due to their enormous size and complexity. A better understanding of the global structure of the network of neurons in the brain could, for example, help understand the spread of the tau protein through the brain, causing Alzheimer's disease [28]. We do know that many real-world networks exhibit the following three properties:

(1) The small-world property: distances within a network are logarithmic or double logarithmic (ultrasmall-world) [36]. Examples include social networks [7, 70, 74], food webs [66], electric power grids and movie actor networks [75].

(2) The scale-free property: the number of neighbours or connections of vertices in the network statistically follows a power-law distribution, i.e.  $\mathbb{P}(\text{degree} > k) \approx k^{-(\tau-1)}$ , for some  $\tau > 1$  [4, 5, 68]. Examples include the World Wide Web and the Internet, [40], global cargo ship movements [59], and food webs [66].

---

*Date:* April 26, 2019.

*2010 Mathematics Subject Classification.* Primary: 60C05, 05C80, 90B15.

*Key words and phrases.* Spatial network models, hyperbolic random graphs, scale-free property, small world property, typical distances, first passage percolation.

(3) High clustering. Local clustering is the formation of locally concentrated clusters in a network and it is measured by the average clustering coefficient: the proportion of triangles present at a vertex versus the total number of possible triangles present at said vertex, averaged over all vertices [76]. Many real-life networks show high clustering, e.g. biological networks such as food webs, technological networks as the World Wide Web, and social networks [30, 64, 69, 70, 72].

The study of real-world networks could be further improved by the analysis of network models. In the last three decades, network science has become a significant field of study, both theoretically and experimentally. One of the first models that was proposed for studying real-world networks was the Watts-Strogatz model [76], that has the small-world property and shows clustering, but does not exhibit the just at that time discovered scale-free property, however. Barabási and Albert proposed that power-law degree distributions are caused by a ‘preferential attachment’ mechanism [8]. The Barabási-Albert model, also known as the preferential attachment model, produces power-law degree distributions intrinsically by describing the growth mechanism of the network over time. Static models, on the contrary, capture a ‘snapshot’ of a network at a given time, and generally, their parameters can be tuned extrinsically. The configuration model and variations of inhomogeneous random graph models such as the Norros-Reitu or Chung-Lu models [20, 21, 27, 71], serve as popular null-models for applied studies as well as accommodate rigorous mathematical analysis. The degree distribution and typical distances are extensively studied for these models [24, 27, 26, 49, 50, 57, 71]. Typical distances in these models show the same qualitative behaviour: the phase transition from small-world to ultrasmall-world takes place when the second moment of the degree (configuration model) or weight distribution (inhomogeneous random graphs) goes from being finite to being infinite. This also follows for the preferential attachment model [35] and it is believed that it holds for a very large class of models.

The one property that most of these models do not exhibit, however, is clustering. As the graph size tends to infinity, most models have a vanishing average clustering coefficient. Therefore, among other methods to increase clustering such as the hierarchical configuration model [52], spatial random graph models have been introduced, including the spatial preferential attachment model (SPA) [3], power-law radius geometric random graphs [46, 77], the hyperbolic random graph (HRG) [18, 19, 63], and the models that we study in this paper, scale-free percolation on  $\mathbb{Z}^d$  (SFP) [32], their continuum analogue on  $\mathbb{R}^d$  [34] and the geometric inhomogeneous random graph (GIRG) [23], respectively. The SFP and GIRG can be seen as the spatial counterparts of the Norros-Reitu and the Chung-Lu graph, respectively. Most of these spatial models are equipped with a long-range parameter  $\alpha > 0$  that models how strongly edges are predicted in relation to their distance. For the SPA, HRG and GIRG, the average clustering coefficient is bounded from below (for the SPA this is shown only when  $\tau < 3$ ) [22, 29, 42, 54], they possess power-law degrees [3, 22, 32, 42] and typical distances in the SFP and GIRG behave similar to other non-spatial models in the infinite variance degree regime [22, 32]. In the finite variance degree regime, geometry starts to play a more dominant role and depending on the long-range parameter  $\alpha$ , distances interpolate from small-world to linear distances (partially known for the SFP) [32, 33].

Understanding the network topology lays the foundation for studying random processes on these networks, such as random walks, or information or activity diffusion. These processes can be found in many real-world networks as well, e.g., virus spreading, optimal targeting of advertisements, page-rank and many more. Mathematical models for information diffusion include the contact process, bootstrap percolation and first passage percolation (FPP). Due to the novelty of spatial scale-free models, the mathematical understanding of processes on them is rather limited. Random walks on scale-free percolation was studied in [45], bootstrap percolation on hyperbolic random graphs and on GIRGs in [25, 60], and FPP on SFP in some regimes in [51]. In this paper we intend to add to this knowledge by studying FPP on GIRG and also on SFP, solving an open problem in [51].

First passage percolation stands for allocating an independent and identically distributed (i.i.d.) random edge weight to each edge from some underlying distribution and study how metric balls grow in the random metric induced by the edge-weights. The edge-weights can be seen as transmission times of information, and hence weighted distances correspond to the time it takes for the information to reach one vertex when started from the other vertex. FPP was originally introduced as a method to model the flow of a fluid through a porous medium in a random setting [43], and has become an

important mathematical model in probability theory that has been studied extensively in grid-like graphs, see e.g. [43, 53, 73], and on non-spatial (random) graph models as well such as the complete graph [48, 37, 38, 39, 56], the Erdős-Rényi graph [14], the configuration model [13, 12, 50] and the inhomogeneous random graph model [61]. FPP in the ‘true’ scale-free regime on the configuration model, (i.e. when the asymptotic variance of the degrees is infinite), was studied recently in [2, 9, 10], revealing a dichotomy of distances caused by the dichotomy of ‘explosiveness’ vs ‘conservativeness’ of the underlying branching processes (BP). In particular, the local tree-like structure allows for the use of age-dependent branching processes when analysing FPP. Age-dependent BPs, introduced by Bellman and Harris in [11], are branching processes where each individual lives for an i.i.d. lifetime and upon death it gives birth to an i.i.d. number of offspring. When the mean offspring is infinite, for some lifetime distributions, it is possible that the BP creates infinitely many individuals in finite time. This phenomenon is called *explosion*. For other lifetime distributions this is not the case, in which case the process is called *conservative*. A necessary and sufficient criterion for explosion was given in a recent paper [6].

*Our contribution* is that we show that the weighted distances in GIRGs and SFP in the regime where the degrees have infinite variance, converge in distribution when the edge-weight distribution produces explosive BPs with infinite mean offspring distributions. We further identify the distributional limit. The case when the edge-weight distribution is conservative has been studied in the SFP model in [51]. Based on this, and the result in [2], we expect that a similar result would hold for the GIRG as well. We formulate our main result without the technical details in the following meta-theorem. Let  $F_L$  denote the probability distribution function of the non-negative random variable  $L$ , and equip every edge  $e$  in the GIRG (resp. SFP) graph with an i.i.d. edge-weight  $L_e$ , a copy of  $L$ . The weight of a path in the network is defined as the sum of the edge-weights of the edges in the path. Let  $d_L(u, v)$  denote the weight of the least-weight path, called weighted distance, between two vertices  $u, v$  in the model (see Def. 2.9 below for a precise definition).

**Theorem 1.1** (Meta-Theorem). *Consider GIRG on  $n$  vertices or SFP on  $\mathbb{Z}^d$  with i.i.d. power-law vertex-weights, and i.i.d. edge-weights with probability distribution function  $F_L$ . Let the parameters of the model be so that the degree distribution follows a power-law with exponent  $\tau \in (2, 3)$ , and  $L$  as lifetime forms an explosive age-dependent branching process with infinite mean power-law offspring distributions. Let  $u, v$  be uniformly chosen vertices in GIRG. Then  $d_L(u, v)$  converges in distribution to an a.s. finite random variable, conditioned that  $u, v$  are in the giant component of GIRG.*

*In the SFP, let  $u := 0$ , the origin of  $\mathbb{Z}^d$ , and  $v$  the closest vertex to  $n\mathbf{e}$  for some arbitrary unit vector  $\mathbf{e}$ . Then  $d_L(u, v)$  converges in distribution to an a.s. finite random variable, conditioned that  $u, v$  are in the infinite component of SFP.*

The criterion on the edge-weight distribution  $L$  seems somewhat vague, but it is explicit. In fact, the necessary and sufficient criterion for a BP to be explosive appeared first in [6] in great generality and for power-law offspring distributions it simplifies to an explicitly computable sum only involving the distribution function  $F_L$ , see [62] for a proof. Namely,  $L$  forms an explosive age-dependent BP with power-law offspring distributions if and only if

$$\sum_{k=1}^{\infty} F_L^{(-1)}(e^{-e^k}) < \infty. \quad (1.1)$$

In [23, Theorem 7] it is shown that *hyperbolic random graphs* (HRG) are a special case of GIRGs, i.e. for every set of parameters in a hyperbolic random graph, there is a set of parameters in GIRG that produce graphs that are equal in distribution. This suggest that our result for GIRGs carries through for hyperbolic random graphs, however, it is not obvious that the parameter setting of HRG as a GIRG satisfies the conditions of Theorem 1.1. We devote Section 9 to show that HRGs (when considered as GIRGs) satisfy all the (hidden) conditions in Theorem 1.1, and hence we obtain the following corollary:

**Corollary 1.2.** *The statement of Theorem 1.1 for GIRG remains word-for-word valid for hyperbolic random graphs as well.*

**Notation.** We write rhs and lhs for right-hand side and left-hand side, respectively, wrt for with respect to, rv for random variable, i.i.d. for independent and identically distributed, pdf for probability distribution function. Generally, we write  $F_X$  for the probability distribution function of a rv  $X$ , and  $F_X^{(-1)}$  for its generalised inverse function, defined as  $F_X^{(-1)}(y) := \inf\{t \in \mathbb{R} : F_X(t) \geq y\}$ , and we write  $\mathcal{L}eb$  for the Lebesgue measure. An event  $E$  happens almost surely (a.s.) when  $\mathbb{P}(E) = 1$  and a sequence of events  $(E_n)_{n \in \mathbb{N}}$  holds with high probability (whp) when  $\lim_{n \rightarrow \infty} \mathbb{P}(E_n) = 1$ . For rvs  $(X_n)_{n \in \mathbb{N}}, X, Y$ , we write  $X_n \xrightarrow{d} X$  and  $X \geq^d Y$  when the  $X_n$  converges in distribution to  $X$ , and when  $X$  stochastically dominates  $Y$ , that is, for all  $x \in \mathbb{R} : F_X(x) \leq F_Y(x)$ , respectively. The rvs  $(X_n)_{n \in \mathbb{N}}$  are tight, if for every  $\varepsilon > 0$  there exists a  $K_\varepsilon > 0$  such that  $\mathbb{P}(|X_n| > K_\varepsilon) < \varepsilon$  for all  $n$ . For  $n \in \mathbb{N}$ , let  $[n] := \{1, \dots, n\}$ , and for two vertices  $u, v$ , let  $u \leftrightarrow v$  denote the event that  $u, v$  are connected by an edge  $e = (u, v)$ .  $\|x - y\|$  denotes the Euclidean norm between  $x$  and  $y$ . We denote by  $\lfloor x \rfloor, \lceil x \rceil$  the lower and upper integer part of  $x \in \mathbb{R}$ , respectively, while, when  $x \in \mathbb{R}^d$ ,  $\lfloor x \rfloor, \lceil x \rceil$  denotes taking the upper/lower integer part element-wise. We write  $x \wedge y := \min\{x, y\}$ . We generally denote vertex/edge sets by  $\mathcal{V}, \mathcal{E}$  while  $|\cdot|$  denotes the size of the set.

## 2. MODEL AND RESULTS

We begin with introducing the Geometric Inhomogeneous Random Graph model (GIRG) from [23].

**Definition 2.1** (Geometric Inhomogeneous Random Graph). *Let  $W^{(n)} \geq 1, L \geq 0$  be random variables. Let  $V := [n]$ , and consider a measure space  $(\mathcal{X}, \nu)$  with  $\nu(\mathcal{X}) = 1$ . Assign to each vertex  $i \in [n]$  an i.i.d. position vector  $x_i \in \mathcal{X}$  sampled from the measure  $\nu$ , and a vertex-weight  $W_i^{(n)}$ , an i.i.d. copy of  $W^{(n)}$ . Then, conditioned on  $(x_i, W_i^{(n)})_{i \in [n]}$  edges are present independently and for any  $u, v \in [n]$ ,*

$$\mathbb{P}(u \leftrightarrow v \text{ in GIRG}_{W,L}(n) \mid (x_i, W_i^{(n)})_{i \in [n]}) := g_n(x_u, x_v, (W_i^{(n)})_{i \in [n]}), \quad (2.1)$$

where  $g_n : \mathcal{X} \times \mathcal{X} \times \mathbb{R}_+^n \rightarrow [0, 1]$  measurable. Finally, assign to each present edge  $e$  an edge-length  $L_e$ , an i.i.d. copy of  $L$ . We denote the resulting graph by  $\text{GIRG}_{W,L}(n)$ .

In [23], when  $\mathcal{X} = \mathcal{X}_d := [0, 1]^d$ , with  $d$  standing for the parameter of the dimension, the following bounds were assumed on  $g_n$ . There is a parameter  $\alpha > 1$ , and  $0 < c_1 \leq C_1 \leq 1$ , such that for all  $n$  and all  $u, v \in [n]$ ,

$$c_1 \leq \frac{g_n(x_u, x_v, (W_i^{(n)})_{i \in [n]})}{1 \wedge \|x_u - x_v\|^{-\alpha d} (W_u^{(n)} W_v^{(n)} / \sum_{i=1}^n W_i^{(n)})^\alpha} \leq C_1. \quad (2.2)$$

In this paper, we use a slightly different assumption that captures a larger class of  $g_n$  when  $\mathbb{E}[W^{(n)}] < \infty$ . Let us introduce two functions,  $\bar{g}, \underline{g} : \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow [0, 1]$ , having parameters  $\alpha, \gamma, \underline{a}_1, \bar{a}_1, a_2 \in \mathbb{R}_+$ :

$$\begin{aligned} \bar{g}(x, w_1, w_2) &:= 1 \wedge \bar{a}_1 \|x\|^{-\alpha d} (w_1 w_2)^\alpha, \\ \underline{g}(x, w_1, w_2) &:= e^{-a_2((\log w_1)^\gamma + (\log w_2)^\gamma)} \wedge \underline{a}_1 \|x\|^{-\alpha d} (w_1 w_2)^\alpha, \end{aligned} \quad (2.3)$$

**Assumption 2.2.** *There exist parameters  $\alpha > 1, \gamma < 1, \bar{a}_1, \underline{a}_1, a_2 \in \mathbb{R}_+$  and  $0 < c_1 \leq C_1 < \infty$ , such that for all  $n$  and all  $u, v \in [n]$ ,  $g_n$  in (2.1) satisfies*

$$c_1 \cdot \underline{g}(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}) \leq g_n(x_u, x_v, (W_i^{(n)})_{i \in [n]}) \leq C_1 \cdot \bar{g}(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}). \quad (2.4)$$

Comparing the upper and lower bounds in (2.4) to those in (2.2), by multiplying the spatial difference  $x_u - x_v$  by  $n^{1/d}$  in  $\bar{g}, \underline{g}$ , we have replaced  $\sum_{i=1}^n W_i^{(n)}$  in (2.2) by a constant times  $n$  in (2.4). For large  $n$  this does not make a difference when  $\mathbb{E}[W^{(n)}] < \infty$ , since in this case the sum is asymptotically  $n\mathbb{E}[W^{(n)}]$  by the Law of Large Numbers. A more important change is that we have altered the first argument of the minimum in the lower bound in (2.2). The reason for the extension of the lower bound in (2.2) to the weaker form in (2.4) is that models satisfying (2.4) still show (asymptotically) the same qualitative behaviour as the ones satisfying (2.2), and when applying weight-dependent percolation, it

becomes necessary to allow edge probabilities to satisfy only the lower bound in (2.4) but not (2.2). We discuss this in more detail in Section 6. A similar model is scale-free percolation (SFP), introduced

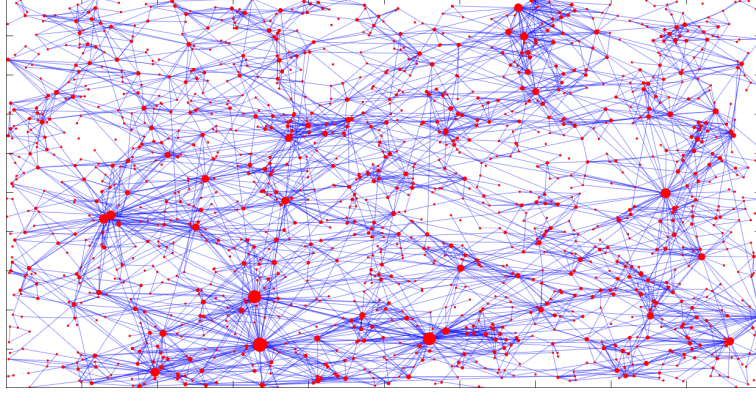


FIGURE 1. A realisation of the GIRG, for  $\alpha = 1.95, \tau = 2.8525, d = 2, n = 1000$ .

by Deijfen, van der Hofstad and Hooghiemstra in [32]. In Section 3 we discuss the relation of GIRG and SFP in more detail. A realisation of (part of) the GIRG model with power-law weights can be seen in Figure 1.

**Definition 2.3** (Scale-free percolation). *Let  $d \geq 1$  be an integer and  $W \geq 1$  a random variable. We assign to each vertex  $i \in \mathbb{Z}^d$  a random weight  $W_i$ , an i.i.d. copy of  $W$ . For a parameter  $\tilde{\alpha} > 0$  and a percolation parameter  $\lambda > 0$ , conditioned on  $(W_i)_{i \in \mathbb{Z}^d}$ , we connect any two non-nearest neighbours  $u, v \in \mathbb{Z}^d$  independently with probability*

$$\mathbb{P}(u \leftrightarrow v \mid \|u - v\| > 1, (W_i)_{i \in \mathbb{Z}^d}) = 1 - \exp \left\{ -\lambda \|u - v\|^{-\tilde{\alpha}} W_u W_v \right\}. \quad (2.5)$$

*Vertices with distance at most one are connected with probability 1. Finally, we assign to each edge  $e$  an edge-length  $L_e$ , an i.i.d. copy of  $L$ . We denote the resulting random graph by  $\text{SFP}_{W,L}$ .*

*Vertex-weight distribution.* To be able to produce power-law degree distributions, in both the GIRG and SFP, it is generally interesting to study the case when the weight variables asymptotically follow a power-law distribution. In the literature, (see e.g. Chung-Lu, or Norros-Reitu model), generally the same vertex-weight distribution  $W^{(n)} \equiv W$  is assumed for all values  $n \geq 1$ . For our results to carry through to hyperbolic random graphs it is necessary to allow  $n$ -dependent weight distributions that converge to some limiting distribution. Hence we pose the following assumption on the weight distributions  $W^{(n)}$ . We say that a function varies slowly at infinity when  $\lim_{x \rightarrow \infty} \ell(cx)/\ell(x) = 1$  for all  $c > 0$ .

**Assumption 2.4** (Power-law limiting weights). *The vertex-weight distributions  $(W^{(n)})_{n \geq 1}$  satisfy the following:*

- (a)  $W^{(n)} \geq 1$  a.s. for all  $n \geq 1$ ,
- (b) For all  $n \geq 1$  there exists an  $M_n \in \mathbb{R}_+$  with the property that  $\mathbb{P}(W^{(n)} > M_n) = o(n^{-1})$  and that for all  $x \in [1, M_n]$ ,

$$\mathbb{P}(W^{(n)} > x) = x^{-(\tau-1)} \ell^{(n)}(x) \quad \text{and} \quad \underline{\ell}(x) \leq \ell^{(n)}(x) \leq \bar{\ell}(x), \quad (2.6)$$

*for some functions  $\underline{\ell}(x), \bar{\ell}(x)$  that vary slowly at infinity.*

- (c) *There exists a distribution  $W$  and a function  $\ell$  that varies slowly at infinity, such that*

$$\mathbb{P}(W > x) = \ell(x) x^{-(\tau-1)}, \quad (2.7)$$

*and  $W^{(n)} \xrightarrow{d} W$ .*



By possibly adjusting  $g_n$ , we have assumed that  $\mathbb{P}(W^{(n)} \geq 1) = 1$  to capture all  $W^{(n)}$  with support separated from 0. Note also that we do not require that  $\ell^{(n)}(x)$  is slowly varying, it is enough if its limit  $\ell(x)$  has this property. The reason for this rather weak assumption is that it allows for slightly truncated power-law distributions, and it is also necessary for HRG: there,  $\ell^{(n)}(x)$  contains a term  $(x/n)^\beta$  for some  $\beta > 0$ , so  $\ell^{(n)}(x)$  in itself is not slowly varying, only its limit, when this term vanishes as  $n \rightarrow \infty$ .

Recall that a joint distribution  $(\widehat{W}^{(n)}, \widehat{W})$  on  $\mathbb{R} \times \mathbb{R}$  is a coupling of  $W^{(n)}$  and  $W$ , if its first and second marginals have the law of  $W^{(n)}$  and  $W$ , respectively. The coupling error is defined as  $\mathbb{P}(\widehat{W}^{(n)} \neq \widehat{W})$ . Recall that the total variation distance equals

$$d_{\text{TV}}(W^{(n)}, W) := \sup_{A \subset \mathbb{R}} |\mathbb{P}(W^{(n)} \in A) - \mathbb{P}(W \in A)| = \inf_{(\widehat{W}^{(n)}, \widehat{W}) \text{ coupling of } W^{(n)}, W} \mathbb{P}(\widehat{W}^{(n)} \neq \widehat{W}). \quad (2.8)$$

the total variation distance between  $W^{(n)}$  and  $W$ . By Assumption 2.4,  $d_{\text{TV}}(W^{(n)}, W) \rightarrow 0$  as  $n \rightarrow \infty$ . Further, whenever  $d_{\text{TV}}(W^{(n)}, W) \rightarrow 0$ , (by definition of the infimum, see also [58, Theorem 4.2]) it is true that there exists a sequence of *couplings* of the rvs  $(W^{(n)})_{n \geq 1}, W$  such that the coupling error tends zero. Let us thus take such a sequence of couplings  $(\widehat{W}^{(n)}, \widehat{W})$  and define

$$\epsilon_{\text{TV}}(n) := \mathbb{P}(\widehat{W}^{(n)} \neq \widehat{W}) \rightarrow 0. \quad (2.9)$$

**Results.** Theorems 2.12 and 2.14 below, our main results, are the precise versions of Theorem 1.1 above. Before formulating the results we introduce some extra assumptions, and models that help us state the distributional limits in our theorems. Let us express

$$g_n(x_u, x_v, (W_i^{(n)})_{i \in [n]}) =: \tilde{g}_n(x_u, x_v, W_u^{(n)}, W_v^{(n)}, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}), \quad (2.10)$$

i.e. the notation  $\tilde{g}_n$  emphasises the weights of the vertices  $u, v$  in  $g_n$  in (2.1). Heuristically, the next assumption ensures that  $g_n$ , the edge-connection probability function, converges to some limiting function  $h$  that only depends on the spatial distance between the two vertices and on their weights.

**Assumption 2.5** (Limiting connection probabilities exists). *Set  $(\mathcal{X}, \nu) := ([-1/2, 1/2]^d, \mathcal{L}eb)$  for some integer  $d \geq 1$ . We assume that*

- (a) *there exists an event  $\mathcal{L}_n$  measurable wrt to the  $\sigma$ -algebra generated by the weights  $(W_i^{(n)})_{i \in [n]}$  that satisfies  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{L}_n) = 1$ ,*
- (b) *there exist intervals  $I_\Delta(n) \subseteq \mathbb{R}_+$ ,  $I_w(n) \subseteq [1, \infty)$  and a sequence  $\epsilon(n) \in \mathbb{R}_+$ , such that  $I_\Delta(n) \rightarrow (0, \infty)$ ,  $I_w(n) \rightarrow [1, \infty)$ ,  $\epsilon(n) \rightarrow 0$  as  $n \rightarrow \infty$ ,*
- (c) *there exists a function  $h : \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow [0, 1]$ ,*

*such that whenever the (fixed) triplet  $(\Delta, w_u, w_v) \in \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+$  satisfies that  $\|\Delta\| \in I_\Delta(n)$ ,  $w_u, w_v \in I_w(n)$ , on the event  $\mathcal{L}_n$ ,  $g_n$  in (2.1) satisfies for  $\nu$ -almost every  $x \in \mathcal{X}$ , all  $u, v \in [n]$  that*

$$\frac{|\tilde{g}_n(x, x + \Delta/n^{1/d}, w_u, w_v, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}) - h(\Delta, w_u, w_v)|}{h(\Delta, w_u, w_v)} \leq \epsilon(n). \quad (2.11)$$

The function  $h$  is the limit of the connection probability function  $g_n$ , when the two vertices involved are distance  $\Delta/n^{1/d}$  away from each other. Since  $\sum_{i=1}^n W_i^{(n)} = \Theta(n)$ , the scale  $\Theta(n^{1/d})$  is the scale for the distance between two vertices when  $g_n$  is of constant order in (2.2). The function  $h$  is important since it captures the limiting connection probabilities, that allows us to extend the model to the whole  $\mathbb{R}^d$ . The arguments of  $h$  represent the distance between two vertices ( $\Delta/n^{1/d}$ ) and their weights, respectively. The need for the event  $\mathcal{L}_n$  is coming from the randomness of the  $(W_i^{(n)})_{i \leq n}$ , a typical choice for  $\mathcal{L}_n$  is e.g.  $\mathcal{L}_n := \{\sum_{i \leq n} W_i^{(n)} - n\mathbb{E}[W^{(n)}] \in (-n^\beta, n^\beta)\}$  for some  $\beta < 1$ .

Importantly, Assumption 2.5 implies that the connection probabilities converge whp to the function  $h(\Delta, w_u, w_v)$ , and this limiting connection function only depends on the vertex-weights and the spatial distance of the vertices involved. Importantly, the limiting probability is *translation invariant*. The

condition requires that in the range where the connection probabilities are not vanishing (i.e. the spatial difference between the two vertices is of order  $n^{-1/d}$ ), the relative error between the connection probability and its limit is uniformly bounded by some function  $\epsilon(n)$ , on intervals that tend to the full possible range. This condition is satisfied when for instance  $g_n$  equals either the upper or the lower bound in (2.2) or in (2.4) and the weights are i.i.d. with  $\mathbb{E}[W^{(n)}] = \mathbb{E}[W] < \infty$ . The bound (2.11) with  $\epsilon(n) = n^{-\kappa(\tau)}$ ,  $I_\Delta(n) = (0, \infty)$ ,  $I_w(n) = [1, \infty)$  for an appropriate  $\kappa(\tau) > 0$  can be shown using that the fluctuations of  $\sum_{i \leq n} W_i^{(n)}$  around  $n\mathbb{E}[W^{(n)}]$  are of order  $n^{1/2}$  whenever  $\tau > 3$  and are  $n^{1/(\tau-1)}$  when  $\tau \in (2, 3)$ , and that for any two positive numbers  $a, b$ ,  $|(1 \wedge a) - (1 \wedge b)| \leq 1 \wedge |a - b|$ . Hyperbolic random graphs satisfy this assumption with the choices  $\epsilon(n) = n^{-1}$  and  $I_\Delta(n) = [n^{-\beta(\alpha)}, n^{\beta(\alpha)}]$ ,  $I_w(n) = [1, n^{\beta(\alpha)}]$  for some  $\beta(\alpha)$ , that we show in Section 9.

**Claim 2.6.** *Suppose  $g_n$  satisfies Assumption 2.2, with some  $\alpha, \gamma$  and Assumption 2.5. Then with the same  $\alpha, \gamma$  and for some  $\underline{c}_1, \overline{C}_1$  with  $0 < \underline{c}_1 \leq c_1 \leq C_1 \leq \overline{C}_1 < \infty$ , the limit  $h$  in (2.11) satisfies*

$$\underline{c}_1 \underline{g}(\Delta, w_u, w_v) \leq h(\Delta, w_u, w_v) \leq \overline{C}_1 \overline{g}(\Delta, w_u, w_v). \quad (2.12)$$

*Proof.* Since  $I_\Delta(n) \rightarrow (0, \infty)$ ,  $I_w(n) \rightarrow [1, \infty)$ , using (2.4), rearranging (2.11) yields that

$$\begin{aligned} h(\Delta, w_u, w_v) &\leq (1 - \epsilon(n))^{-1} g_n(x, x + \Delta/n^{1/d}, (W_i^{(n)})_{i \in [n]}) \leq (1 - \epsilon(n))^{-1} C_1 \overline{g}(\Delta, w_u, w_v), \\ h(\Delta, w_u, w_v) &\geq (1 + \epsilon(n))^{-1} \underline{c}_1 \underline{g}(\Delta, w_u, w_v), \end{aligned} \quad (2.13)$$

for all sufficiently large  $n$ , thus the result follows by adjusting the constants  $c_1, C_1$ .  $\square$

To be able to define the distributional limit of weighted distances in GIRG, we need to define an infinite random graph, that we think of as the *extension of  $\text{GIRG}_{W,L}(n)$  to  $\mathbb{R}^d$* . This model is a generalisation of [34], and it can be shown that the continuum percolation model in [41], the (local weak) limit of hyperbolic random graphs, can be transformed to this model with specific parameters that we identify in Section 9 below.

**Definition 2.7** (Extended GIRG). *Let  $h : \mathbb{R}^d \times \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow [0, 1]$  be a function,  $W \geq 1, L \geq 0$  random variables and  $\lambda > 0$ . We define the infinite random graph model  $\text{EGIRG}_{W,L}(\lambda)$  as follows. Let  $\mathcal{V}_\lambda$  be a homogeneous Poisson Point Process (PPP) on  $\mathbb{R}^d$  with intensity  $\lambda$ , forming the positions of vertices. For each  $x \in \mathcal{V}_\lambda$  draw a random weight  $W_x$ , an i.i.d. non-negative rv with pdf  $F_W$ . Then, conditioned on  $(x, W_x)_{x \in \mathcal{V}_\lambda}$ , edges are present independently with probability*

$$\mathbb{P}(y \leftrightarrow z \text{ in } \text{EGIRG}_{W,L}(\lambda) \mid (x, W_x)_{x \in \mathcal{V}_\lambda}) := h(y - z, W_y, W_z). \quad (2.14)$$

*Finally, we assign to each present edge  $e$  an edge-length  $L_e$ , an i.i.d. copy of  $L$ . We write  $(\mathcal{V}_\lambda, \mathcal{E}_\lambda)$  for the vertex and edge set of the resulting graph, that we denote by  $\text{EGIRG}_{W,L}(\lambda)$ .*

Note that in the extension it is necessary to use the limiting function  $h$  instead of  $g_n$ , since  $g_n$  is not defined outside  $\mathcal{X}_d$ . We also use the limiting weight distribution  $W$  instead of  $W^{(n)}$  for the weights of vertices.

To be able to connect  $\text{GIRG}_{W,L}(n)$  to the extended infinite model  $\text{EGIRG}_{W,L}(\lambda)$ , we ‘blow-up’  $\text{GIRG}_{W,L}(n)$  on  $(\mathcal{X}, \nu) = ([-1/2, 1/2]^d, \mathcal{L}eb)$ , by a factor  $n^{1/d}$  so that the average density of vertices becomes 1. This is the next model we define.

**Definition 2.8** (Blown-up  $\text{GIRG}_{W,L}(n)$ ). *Let  $W^{(n)} \geq 1, L \geq 0$ , and  $g_n$  be from Definition 2.1. Let  $(\tilde{\mathcal{X}}_d(n), \nu_n) := ([-n^{1/d}/2, n^{1/d}/2]^d, \mathcal{L}eb)$ . Assign to each vertex  $i \in [n]$  an i.i.d. position vector  $x_i \in \tilde{\mathcal{X}}_d(n)$  sampled from the measure  $\mathcal{L}eb$ , and a vertex-weight  $W_i^{(n)}$ , an i.i.d. copy of  $W^{(n)}$ . Then, conditioned on  $(x_i, W_i^{(n)})_{i \in [n]}$  edges are present independently and for any  $u, v \in [n]$ ,*

$$\begin{aligned} \mathbb{P}(y \leftrightarrow z \text{ in } \text{BGIRG}_{W,L}(n) \mid (x_i, W_i^{(n)})_{i \in [n]}) &:= g_n(x_u/n^{1/d}, x_v/n^{1/d}, (W_i^{(n)})_{i \in [n]}) \\ &=: g_n^B(x_u, x_v, (W_i^{(n)})_{i \in [n]}) \end{aligned} \quad (2.15)$$

*Finally, assign to each present edge  $e$  an edge-length  $L_e$ , an i.i.d. copy of  $L$ . We write  $\mathcal{V}_B(n), \mathcal{E}_B(n)$  for the vertex and edge-set of the resulting graphs, that we denote by  $\text{BGIRG}_{W,L}(n)$ .*



Naturally,  $\text{BGIRG}_{W,L}(n) \stackrel{d}{=} \text{GIRG}_{W,L}(n)$  since it can be obtained as a deterministic transformation of the model defined in Definition 2.1, namely, mapping every vertex  $x_u \in [-1/2, 1/2]^d$  to  $n^{1/d}x_u$  to obtain a vertex in  $\text{BGIRG}_{W,L}(n)$ . The advantage of working with the blown-up version is that it naturally yields a limiting infinite graph on  $\mathbb{R}^d$ , which is  $\text{EGIRG}_{W,L}(1)$ . By (2.15) and (2.4), Assumption 2.2 for  $\text{BGIRG}_{W,L}(n)$  reduces to

$$c_1 g(\|x_u - x_v\|, W_u^{(n)}, W_v^{(n)}) \leq g_n^B(x_u, x_v, (W_i^{(n)})_{i \in [n]}) \leq C_1 \bar{g}(\|x_u - x_v\|, W_u^{(n)}, W_v^{(n)}). \quad (2.16)$$

Next we define weighted and graph distance, and introduce notation for corresponding metric balls and their boundaries, and then define the *explosion time* of the origin in  $\text{EGIRG}_{W,L}(\lambda)$ .

**Definition 2.9** (Distances and metric balls). *For a path  $\pi$ , we define its length and  $L$ -length as  $|\pi| = \sum_{e \in \pi} 1$ , and  $|\pi|_L := \sum_{e \in \pi} L_e$ , respectively. Let  $\Omega_{u,v} := \{\pi : \pi \text{ is a path from } u \text{ to } v\}$ . For two sets  $A, B \subseteq V$ , we define the  $L$ -distance and the graph distance as*

$$d_L(A, B) := \inf_{u \in A} \inf_{v \in B} \inf_{\pi \in \Omega_{u,v}} |\pi|_L, \quad d_G(A, B) := \inf_{u \in A} \inf_{v \in B} \inf_{\pi \in \Omega_{u,v}} |\pi|. \quad (2.17)$$

with  $d_q(A, B) = 0$  when  $A \cap B \neq \emptyset$  and  $d_q(A, B) = \infty$  when there is no path from  $A$  to  $B$  for  $q \in \{L, G\}$ . For Euclidean distance, graph distance and  $L$ -distance, respectively, let

$$\begin{aligned} B^2(u, r) &:= \{v \in V : \|x_u - x_v\| \leq r\}, \\ B^G(u, r) &:= \{v \in V : d_G(u, v) \leq r\}, \\ B^L(u, r) &:= \{v \in V : d_L(u, v) \leq r\}, \end{aligned} \quad (2.18)$$

denote the balls of radius  $r$  with respect to these distances, where  $x_u, x_v$  denote the possibly random positions of vertices  $u$  and  $v$ , respectively. For an integer  $k \geq 1$  we write  $\partial B^G(u, k) := B^G(u, k) \setminus B^G(u, k-1)$  as the set of vertices that are at graph distance  $k$  from  $u$ , and  $B^{L,G}(u, k)$  and  $\partial B^{L,G}(u, k)$  for those vertices  $v$  where the path realising  $d_L(u, v)$  uses at most and precisely  $k$  edges, respectively. We add a respective subscript  $\lambda$  and  $n$  to these quantities when the underlying graph is  $\text{EGIRG}_{W,L}(\lambda)$  and  $\text{BGIRG}_{W,L}(n)$ , defined in Definitions 2.7, 2.8, respectively.

**Definition 2.10** (Explosion time). *Consider  $\text{EGIRG}_{W,L}(\lambda)$  as in Definition 2.7, and let  $v \in \mathcal{V}_\lambda$ . Let  $\tau_\lambda^E(v, k) := \inf\{t : |B^L(v, t)| = k \text{ in } \text{EGIRG}_{W,L}(\lambda)\}$ . Let the explosion time of vertex  $v$  be defined as the (possibly infinite) almost sure limit:*

$$Y_\lambda^E(v) := \lim_{k \rightarrow \infty} \tau_\lambda^E(v, k). \quad (2.19)$$

The explosion time of the origin,  $Y_\lambda^E(0)$ , is defined analogously when  $\mathcal{V}_\lambda$  is conditioned on having a vertex at  $0 \in \mathbb{R}^d$ . We call  $\text{EGIRG}_{W,L}(\lambda)$  explosive if  $\mathbb{P}(Y_\lambda^E(0) < \infty) > 0$ . We call any infinite path  $\pi$  with  $|\pi|_L < \infty$  an explosive path.

$B^L(v, t)$ , defined in (2.18), stands for the vertices that are available within  $L$ -distance  $t$ . Hence,  $\tau_\lambda^E(v, k)$  is the smallest radius  $t$ , in terms of the  $L$ -distance, such that the metric ball  $B^L(v, t)$  around vertex  $v \in \mathcal{V}_\lambda$  contains  $k$  vertices<sup>1</sup>. Then,  $Y_\lambda^E(v)$  is the almost sure limit of  $\tau_\lambda^E(v, k)$  as  $k$  tends to infinity. In most common scenarios considering infinite graphs, this limit is infinite.  $Y_\lambda^E(v)$  becomes finite when it is possible to reach infinitely many vertices within finite  $L$ -distance from  $v$ . In particular, if there is an infinite path of vertices  $\pi = (\pi_0 = v, \pi_1, \pi_2, \dots)$  with total  $L$ -length  $|\pi|_L < \infty$ , then  $Y_\lambda^E(v)$  is finite. In reverse, if the graph is locally finite, that is, all degrees are less than infinity, then one can show (e.g., by arguing by contradiction, as done below in (3.19)) that  $Y_\lambda^E(v) < \infty$  implies the existence of a path  $\pi$  emanating from  $v$  with finite total  $L$ -length. In particular, one can define

$$\pi_{\text{opt}}^1(v) := \arg \inf_{\pi : \pi_0 = v, |\pi|_L = \infty} \{|\pi|_L\} \quad (2.20)$$

be the ‘shortest’ explosive path to infinity from  $u$  in  $\text{EGIRG}_{W,L}(1)$ , that is, the path that realises the shortest  $L$ -length of all paths with infinitely many edges from  $u$ . In case this path is not unique (which

<sup>1</sup>When considering  $t$  as time, this is the time that  $B^L(v, \cdot)$  “swallows” the  $k$ th vertex as one increases  $t$

may happen if  $L$  is not absolutely continuous) then any path realising the infimum can be chosen. It can generally be shown that

$$|\pi_{\text{opt}}^1(u)|_L = Y_1^E(u). \quad (2.21)$$

This phenomenon is called explosion, and it was first observed in branching processes. See Harris [44] for one of the first conditions derived for a continuous-time branching processes to be explosive.

Our first result, analogous to [51, Theorem 1.1] states that  $\text{EGIRG}_{W,L}(\lambda)$  is not explosive when  $\mathbb{E}[W^2] < \infty$ , and characterise the set of edge-length distributions  $L$  that give explosivity when  $\tau \in (2, 3)$  in Assumption 2.4. We denote the distribution function of  $L$  by  $F_L$ .

**Theorem 2.11.** *Consider  $\text{EGIRG}_{W,L}(\lambda)$  in Definition 2.7. Then,  $\text{EGIRG}_{W,L}(\lambda)$  is not explosive whenever  $\mathbb{E}[W^2] < \infty$ . That is,  $Y_\lambda^E(0) = \infty$  almost surely when  $\mathbb{E}[W^2] < \infty$ . Let*

$$\mathbf{I}(L) := \int_1^\infty F_L^{(-1)}(\exp\{-e^y\}) dy. \quad (2.22)$$

*When  $W$  satisfies (2.7) with  $\tau \in (2, 3)$ , then  $\text{EGIRG}_{W,L}(\lambda)$  is explosive if and only if  $\mathbf{I}(L) < \infty$ .*

We are ready to state the two main theorems. By Bringmann *et al.* in [22, Theorem 2.2], whp there is a unique linear-sized giant component  $\mathcal{C}_{\max}$  in  $\text{GIRG}_{W,L}(n)$  when  $W$  satisfies (2.7) with  $\tau \in (2, 3)$ .

**Theorem 2.12** (Distances in GIRG with explosive edge-lengths). *Consider  $\text{GIRG}_{W,L}(n)$ , satisfying Assumptions 2.2, 2.4, 2.5 with parameters  $\tau \in (2, 3)$ ,  $\alpha > 1$ ,  $d \geq 1$ , and assume that  $\mathbf{I}(L) < \infty$ . Let  $v_n^1, v_n^2 \in \mathcal{C}_{\max}$  be two vertices chosen uniformly at random. Then, as  $n$  tends to infinity,*

$$d_L(v_n^1, v_n^2) \xrightarrow{d} Y^{(1)} + Y^{(2)}, \quad (2.23)$$

*where  $Y^{(1)}, Y^{(2)}$  are i.i.d. copies of  $Y_1^E(0)$  from Definition 2.10 conditioned to be finite.*

**Corollary 2.13.** *If all assumptions in Theorem 2.12 hold except Assumption 2.5, then  $(d_L(v_n^1, v_n^2))_{n \geq 1}$  forms a tight sequence of random variables.*

*Proof of Corollary 2.13 subject to Theorem 2.12.* Consider a model  $\underline{\text{GIRG}}_{W,L}(n)$  where  $c_1 g(\cdot)$  is used instead of  $g_n(\cdot)$  on the same vertex positions and vertex-weights  $(x_i, W_i^{(n)})_{i \in [n]}$ . Under Assumption 2.2, a coupling can be constructed where the edge sets  $\mathcal{E}(\underline{\text{GIRG}}_{W,L}(n)) \subseteq \mathcal{E}(\text{GIRG}_{W,L}(n))$  a.s. under the coupling measure. Note that  $c_1 g(\cdot)$  satisfies Assumption 2.5, and so (2.23) is valid for  $\underline{\text{GIRG}}_{W,L}(n)$  and thus Theorem 2.12 holds for  $\underline{d}_L(v_n^1, v_n^2)$ , the  $L$ -distance in  $\underline{\text{GIRG}}_{W,L}(n)$ . The proof of tightness is finished by noting that  $d_L(v_n^1, v_n^2) \leq \underline{d}_L(v_n^1, v_n^2)$  a.s. under the coupling.  $\square$

A similar theorem holds for the SFP model. One direction of this theorem has already appeared in [51, Theorem 1.7], where it was conjectured that its other direction also holds. For a vertex  $x \in \mathbb{Z}^d$  define  $\tau^S(x, k) := \inf\{t : |B^L(x, t)| = k \text{ in } \text{SFP}_{W,L}\}$  and  $Y^S(x) := \lim_{k \rightarrow \infty} \tau^S(x, k)$  as the explosion time of vertex  $x$  in SFP.

**Theorem 2.14** (Distances in SFP with explosive edge-lengths). *Consider  $\text{SFP}_{W,L}$ , with  $W$  satisfying (2.7) with parameter  $\tilde{\tau} > 1$ ,  $\tilde{\alpha} > d$  and  $\gamma = \tilde{\alpha}(\tilde{\tau} - 1)/d \in (1, 2)$  and edge-length distribution  $F_L$  with  $\mathbf{I}(L) < \infty$ . Fix a unit vector  $\underline{e}$ . Then, as  $m \rightarrow \infty$ ,*

$$d_L(0, \lfloor m \underline{e} \rfloor) - (Y^S(0) + Y^S(\lfloor m \underline{e} \rfloor)) \xrightarrow{\mathbb{P}} 0, \quad (2.24)$$

*and*

$$(Y^S(0), Y^S(\lfloor m \underline{e} \rfloor)) \xrightarrow{d} (Y_{(1)}^S, Y_{(2)}^S), \quad (2.25)$$

*where  $Y_{(1)}^S, Y_{(2)}^S$  are two independent copies of the explosion time of the origin  $Y^S(0)$  in  $\text{SFP}_{W,L}$ .*

**2.1. Discussion and open problems.** Understanding the behaviour of distances and weighted distances on spatial network models is a problem that is still widely open, when the graph has a power-law degree distribution. Graph distances are somewhat better understood, at least in the *infinite variance degree* regime, where a giant/infinite component exists whp and typical distances are doubly-logarithmic. For hyperbolic random graphs (HRG), typical distances and the diameter were recently studied in [1] and [67], respectively, and in GIRGs in [22]. For scale-free percolation and its continuum-space analogue, typical distances were studied in [32, 51] and [33]. In the first two papers, doubly-logarithmic distances were established in the infinite variance degree regime, while lower bounds on graph distances were given in other regimes.

Typical distances in the *finite variance regime* remain open - even the existence of a giant component in this case is open for GIRGs, and was studied for HRG in [18, 41]. We conjecture that this result would carry through for GIRGs as well, i.e. a unique linear sized giant component exists when  $\tau > 3$  and the edge-density is sufficiently high. An indication for this is that the limiting continuum percolation model identified in [41] is a special case of the limit of GIRGs, as we show in Section 9. For scale-free percolation (SFP), the phase transition for the existence of a unique infinite component for sufficiently high edge-densities was shown in [32] and for the continuum analogue in [34]. In [33], it is shown that distances grow linearly with the norm of vertices involved, when  $\tilde{\alpha} > 2d$  and degrees have finite variance. The  $\tilde{\alpha} \in (d, 2d)$  case remains open, and they are believed to grow poly-logarithmically with an exponent depending on  $\tilde{\alpha}, d, \tau$  based on the analogous results about *long-range percolation* [16, 17].

Scale-free percolation is the only scale-free spatial model for which results on *first passage percolation* are known [51]. In this paper we extend this knowledge by studying first passage percolation in the infinite variance regime on GIRG, HRG, and SFP, when the edge-length distribution produces explosive branching processes. The non-explosive case was treated in [51] for SFP, and we expect that the results there would carry over GIRG and HRG as well, in an  $n$ -dependent form that is e.g. stated for the configuration model in [2].

Thus, we see the following as main *open questions* regarding (weighted) distances in spatial scale-free network models: Is there universality in these models *beyond* the infinite variance degree regime? More specifically, in the finite variance degree regime: Can we give a description of the interpolation between log-log and linear distances in SFP? Is it similar to long-range percolation? Can we observe the same phenomenon in GIRG and HRG, in terms of the long-range parameter  $\alpha$  and  $T_H$ , respectively? For instance, do distances become linear when the long-range parameter is sufficiently high? Does the Spatial Preferred Attachment model behave similarly to these models?

*Comparison of GIRG and SFP.* The GIRG on  $(\mathcal{X}_d, \nu) = ([0, 1]^d, \mathcal{L}^d)$  uses  $n$  vertices distributed uniformly at random. When blown up to  $[-n^{1/d}/2, n^{1/d}/2]^d$ , the density of the points becomes unit, and, in case the connection probabilities behave somewhat regularly, the model can be extended to  $\mathbb{R}^d$  as we did by introducing the  $\text{EGIRG}_{W,L}(\lambda)$  model. The main difference between GIRG and SFP is thus that their vertex set differs. In the SFP the vertex set is  $\mathbb{Z}^d$ , whereas it is a PPP on  $\mathbb{R}^d$  in  $\text{EGIRG}_{W,L}(\lambda)$ . We conjecture that the local weak limit of  $\text{GIRG}_{W,L}(n)$  satisfying Assumption 2.5 exists and equals  $\text{EGIRG}_{W,L}(1)$ . Beyond this, assuming that GIRG satisfies Assumption 2.2, the two models are differently parametrised. It can be easily shown that the connection probability of the SFP model in (2.5) satisfies the bounds in (2.2), when setting  $\alpha = \tilde{\alpha}/d$ ,  $\tau - 1 = (\tilde{\tau} - 1)\alpha$  and implying that  $\gamma = \tilde{\alpha}(\tilde{\tau} - 1)/d = \tau - 1$ . It follows that where  $\gamma$  is the relevant parameter in the SFP model,  $\tau - 1$  is the relevant parameter in the GIRG model. The relation  $\alpha = \tilde{\alpha}/d$  directly shows why we work with  $\alpha > 1$  for GIRGs: For the SFP  $\tilde{\alpha} > d$  ensures that degrees are finite a.s., while  $\tilde{\alpha} < d$  implies a.s. infinite degrees [32, Theorem 2.1]. This translates to  $\alpha > 1$  for the GIRG model.

**2.2. Overview of the proof.** We devote this section to discussing the strategy of the proof of Theorem 2.12. We start by mapping the underlying space,  $\mathcal{X}_d := [-1/2, 1/2]^d$ , and the vertex locations to  $\tilde{\mathcal{X}}_d := [-n^{1/d}/2, n^{1/d}/2]^d$  using the blow-up method described in Definition 2.8. So, instead of  $\text{GIRG}_{W,L}(n)$ , we work with the equivalent  $\text{BGIRG}_{W,L}(n)$  that has the advantage that the vertex-density stays constant as  $n$  increases. Under Assumptions 2.5, we expect that this model is close to

EGIRG $_{W,L}(1)$  restricted to  $\tilde{\mathcal{X}}_d(n)$ , and thus the shortest paths leaving a large neighbourhood of  $u$  and  $v$  have length that are distributed as the explosion time of those vertices in EGIRG $_{W,L}(1)$ . Since  $u, v$  are typically ‘far away’, these explosion times become asymptotically independent. We show that this heuristics is indeed valid and also that these explosive rays can be *connected to each other* within  $\tilde{\mathcal{X}}_d(n)$ .

The details are quite tedious. We choose a parameter  $\xi_n \downarrow 0$  such that EGIRG $_{W,L}(1 - \xi_n)$ , (resp. EGIRG $_{W,L}(1 + \xi_n)$ ) has, for all large  $n$ , less (resp. more) vertices than  $n$  in  $\tilde{\mathcal{X}}_d(n)$ , as proved in Claim 3.2. In Claim 3.3 below we construct a coupling such that the vertex sets within  $\tilde{\mathcal{X}}_d(n)$  are subsets of each other, i.e. for all  $n$  large enough,

$$\mathcal{V}_{1-\xi_n} \cap \tilde{\mathcal{X}}_d(n) \subseteq \mathcal{V}_B(n) \subseteq \mathcal{V}_{1+\xi_n} \cap \tilde{\mathcal{X}}_d(n), \quad (2.26)$$

and for the edge-sets,  $\mathcal{E}_{\lambda_1} \subseteq \mathcal{E}_{\lambda_2}$  whenever  $\lambda_1 \leq \lambda_2$ . By Def. 2.7, we use the *limiting* edge probabilities  $h$  and weights  $W$  when determining the edges in EGIRG $_{W,L}(\lambda)$ , while in BGIRG $_{W,L}(n)$  we use  $g_n$  and  $W^{(n)}$ . As a result, we cannot hope that a relation similar to (2.26) holds for the edge set of the three graphs as well. Nevertheless, we construct a coupling in Claim 3.4 below, using the error bound in (2.11) (that hold on the good event  $\mathcal{L}_n$  such that for any set of  $k$  edges  $E = (e_1, \dots, e_k)$  where  $e_i$  connects vertices in  $\mathcal{V}_B(n)$ )

$$\mathbb{P}((E \cap \mathcal{E}_B(n)) \ominus (E \cap \mathcal{E}_{1+\xi_n}) \neq \emptyset) \rightarrow 0, \quad (2.27)$$

where  $\ominus$  denotes symmetric difference, as long as the number of vertices involved is not too large. So, any constructed edge-set in one model is present in the other model as well as long as the corresponding lhs in (2.27) stays small. Using the same edge-length  $L_e$  on edges that are present in both models, we can relate the  $L$ -distances in one model to the other.

*Lower bound on  $d_L(v_n^1, v_n^2)$ .* Recall from Def. 2.9 that we add subscript  $\lambda$  and  $n$  to the metric balls and their boundaries when the underlying model is EGIRG $_{W,L}(\lambda)$  and BGIRG $_{W,L}(n)$ , respectively, and the coupling of the vertex sets in (2.26). Suppose we can find an increasing sequence  $k_n \rightarrow \infty$  such that the graph-distance balls  $B_{1+\xi_n}^G(v_n^1, k_n) \cap B_{1+\xi_n}^G(v_n^2, k_n) = \emptyset$  and that both  $B_{1+\xi_n}^G(v_n^1, k_n), B_{1+\xi_n}^G(v_n^2, k_n) \subset \tilde{\mathcal{X}}_d(n)$ . Then, any path connecting  $v_n^1, v_n^2$  in EGIRG $_{W,L}(1 + \xi_n)$  must intersect  $\partial B_{1+\xi_n}^G(v_n^1, k_n)$  and  $\partial B_{1+\xi_n}^G(v_n^2, k_n)$ , hence,

$$d_L^{1+\xi_n}(v_n^1, v_n^2) \geq d_L^{1+\xi_n}(v_n^1, \partial B_{1+\xi_n}^G(v_n^1, k_n)) + d_L^{1+\xi_n}(v_n^2, \partial B_{1+\xi_n}^G(v_n^2, k_n)), \quad (2.28)$$

Let  $A_k$  be the event that  $B_n^G(v_n^1, k)$  and  $B_{1+\xi_n}^G(v_n^1, k)$  as well as  $B_1^G(v_n^1, k)$  (the graph distance balls within graph distance  $k$ ) coincide as graphs with vertex-weights (i.e. they have identical vertex and edge-sets and vertex-weights) and that  $B_{1+\xi_n}^G(v_n^1, k)$  and  $B_{1+\xi_n}^G(v_n^2, k)$  are disjoint, that is,

$$\begin{aligned} A_k := \bigcap_{q \in \{1,2\}} \Big\{ & \mathcal{V}(B_n^G(v_n^q, k)) = \mathcal{V}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{V}(B_1^G(v_n^q, k)), \\ & \mathcal{E}(B_n^G(v_n^q, k)) = \mathcal{E}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{E}(B_1^G(v_n^q, k)), \\ & \mathcal{W}(B_n^G(v_n^q, k)) = \mathcal{W}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{W}(B_1^G(v_n^q, k)) \Big\} \\ & \bigcap \{B_{1+\xi_n}^G(v_n^1, k) \cap B_{1+\xi_n}^G(v_n^2, k) = \emptyset\}, \end{aligned} \quad (2.29)$$

where  $\mathcal{W}(A)$  denotes the vertex-weights of vertices within a set  $A$ . We show in Proposition 3.6 below that on  $A_{k_n}$ , the shortest path connecting  $v_n^q$  to  $\partial B_{1+\xi_n}^G(v_n^q, k_n)$  is also the shortest in BGIRG $_{W,L}(n)$  and in EGIRG $_{W,L}(1)$  as well. Thus, on  $A_{k_n}$ , the lhs of (2.28) can be switched to  $d_L(v_n^1, v_n^2)$  ( $L$ -distance in BGIRG $_{W,L}(n)$ ), and  $1 + \xi_n$  can be changed to 1 on the rhs. Then, it holds that,

$$d_L(v_n^1, v_n^2) \geq \mathbb{1}_{A_{k_n}}(d_L^1(v_n^1, \partial B_1^G(v_n^1, k_n)) + d_L^1(v_n^2, \partial B_1^G(v_n^2, k_n))), \quad (2.30)$$

To find  $k_n$ , for  $q \in \{1, 2\}$ , we bound the maximum displacement within  $B_{1+\xi_n}^G(v_n^q, k)$  in Proposition 4.4. Namely, we couple the exploration of the neighbourhoods of  $v_n^1, v_n^2$  to two dominating branching random walks (BRW) on the vertices of EGIRG $_{W,L}(1 + \xi_n)$ . That is, the vertex set of  $B_{1+\xi_n}^G(v_n^q, k_n)$  is a subset of the location of individuals in generation at most  $k_n$  in the BRWs. Then, we analyse

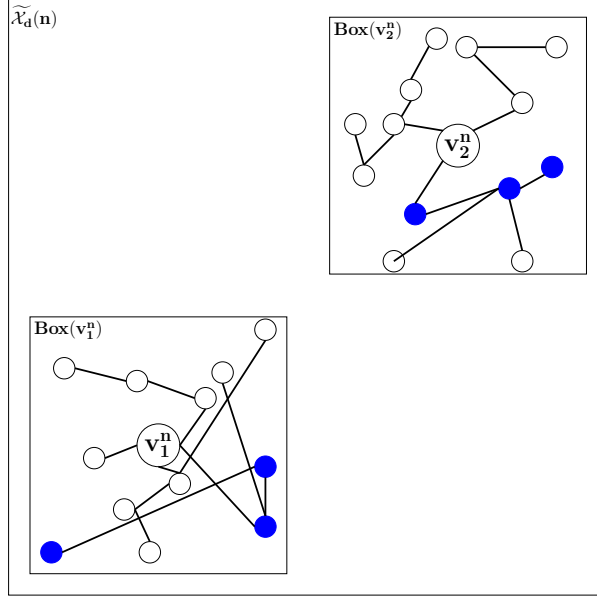


FIGURE 2. The exploration of the neighbourhood of  $v_n^1, v_n^2$  up to graph distance 3, contained in the boxes  $\text{Box}(v_n^1), \text{Box}(v_n^2)$ , with the shortest paths in blue. The idea of the lower bound is: As long as the two boxes are disjoint, and each vertex within graph distance  $k$  from  $v_n^q$  stays within these boxes, the shortest path connecting  $v_n^1, v_n^2$  is longer than the total length of the two shortest paths leading to vertices at graph distance  $k$ . We establish the bound by being able to choose  $k = k_n$  that tends to infinity.

the spatial growth of the BRW. This leads us to show in Proposition 3.6, that there are two spatially disjoint boxes  $\text{Box}(v_n^q), q \in \{1, 2\}$ , and an increasing sequence  $k_n \rightarrow \infty$  such that

$$B_{1+\xi_n}^G(v_n^q, k_n) \subset \cup_{i \leq k_n} \mathcal{G}_i(v_n^q) \subset \text{Box}(v_n^q), \quad (2.31)$$

where  $\mathcal{G}_i(v_n^q)$  denotes the set of generation  $i$  individuals in the BRW started from  $v_n^q$ , and the first containment follows by the coupling between the BRWs and the exploration of the neighbourhoods, see also Figure 2.

This guarantees that the required conditions are satisfied in (2.28). Since the  $B_{1+\xi_n}^G(v_n^q, k_n)$  are defined within disjoint boxes, they are independent and thus the two terms on the rhs of (2.28) are independent. Further, by possibly taking a new sequence  $k_n$  that increases at a slower rate than the previous one, (2.27) ensures that the event  $A_{k_n}$  holds whp in (2.30). We finish the proof by showing that the variables on the rhs of (2.30) tend to two i.i.d. copies of the explosion time of the origin in  $\text{EGIRG}_{W,L}(1)$ , a non-trivial part in itself.

*Upper bound on  $d_L(v_n^1, v_n^2)$ .* For the upper bound, we use the model  $\text{EGIRG}_{W,L}(1 - \xi_n)$ . Under the assumption that both  $v_n^1, v_n^2$  are in the giant component of  $\text{BGIRG}_{W,L}(n)$ , we show in Claim 7.1 that they are whp in the infinite component in the  $\text{EGIRG}_{W,L}(1)$ . Recall from Definition 2.10 that a path  $\pi$  is called explosive if  $|\pi_L| < \infty$ , and that on the event that the explosion time  $Y_1^E(v) < \infty$ , there are explosive paths in the model. Then, recall  $\pi_{\text{opt}}^1(v)$  from (2.20), the ‘shortest’ explosive path to infinity from  $v$  in  $\text{EGIRG}_{W,L}(1)$ . In case this path is not unique (which may happen if  $L$  is not absolutely continuous) then any path realising the infimum can be chosen. The advantage of this path is that  $|\pi_{\text{opt}}^1(v)|_L = Y_1^E(v)$  (see (2.21)) i.e. this path realises the explosion time of  $v$ . Thus, to obtain an upper bound on the  $L$ -distance between  $v_n^1$  and  $v_n^2$ , our goal is to show that the initial segments of the paths  $\pi_{\text{opt}}^1(v_n^1), \pi_{\text{opt}}^1(v_n^2)$  are part of  $\text{BGIRG}_{W,L}(n)$  and that we can *connect* these segments within  $\text{BGIRG}_{W,L}(n)$  with a connecting path of length at most  $\varepsilon$ , say, for arbitrarily small  $\varepsilon > 0$ . The first

part, namely, that the initial segments of these paths are part of  $\text{BGIRG}_{W,L}(n)$ , holds whp since it is a consequence of the event  $A_{k_n}$ : any path in  $\text{EGIRG}_{W,L}(1)$  that leaves  $v_n^q$  and has at most  $k_n$  edges is part of  $\text{BGIRG}_{W,L}(n)$  as well on  $A_{k_n}$ . To be able to connect these segments, we need to find a vertex on  $\pi_{\text{opt}}^1(v_n^1)$  and  $\pi_{\text{opt}}^1(v_n^2)$ , respectively, that has high enough weight.

For a constant  $K$ , define now  $\tilde{v}(K)$  as the first vertex on  $\pi_{\text{opt}}^1(v)$  with weight larger than  $K$  and  $T_K(v) := d_L^1(v, \tilde{v}(K))$ , where  $d_L^1$  denotes the  $L$ -distance in  $\text{EGIRG}_{W,L}(1)$ . Observe that  $T_K(v)$  is the length of the segment between  $v, \tilde{v}(K)$  on the path  $\pi_{\text{opt}}^1(v)$ , otherwise the optimality of  $\pi_{\text{opt}}^1(v)$  would be violated. We show in Corollary 4.2 that  $\tilde{v}(K)$  and  $T_K(v)$  exist: explosion is impossible in the subgraph of  $\text{EGIRG}_{W,L}(\lambda)$  restricted to vertices with weight  $\leq K$ , irrespective of the value of  $\lambda$ . Thus, any infinite path with finite total length must leave this subgraph.

We need to make sure that  $\tilde{v}_n^1(K)$  and  $\tilde{v}_n^2(K)$  are part of  $\text{BGIRG}_{W,L}(n)$ . For this we argue as follows: By translation invariance, for any fixed  $K < \infty$  the spatial distance  $\|u - \tilde{u}(K)\|$  is a copy of the random variable  $\|0 - \tilde{0}(K)\|$ , i.e. a proper random variable, where  $0$  denotes the origin of  $\mathbb{R}^d$ , and  $\tilde{0}(K)$  denotes the first vertex with weight at least  $K$  on the shortest path from  $0$  to infinity in  $\text{EGIRG}_{W,L}(1)$ , conditioned that  $0$  is in the infinite component. Thus, for fixed  $K$ , for  $n$  sufficiently large,  $\tilde{v}_n^q(K) \in \tilde{\mathcal{X}}_d(n)$  will be satisfied with probability tending to 1. Further, we show in Lemma 3.7 that for any fixed  $K$ , the event  $\tilde{A}_{n,K}$  that the vertices and edges of  $\pi_{\text{opt}}^1(v_n^q)$  on the section leading to  $\tilde{v}_n^q(K)$  are present in  $\text{BGIRG}_{W,L}(n)$  as well as in  $\text{EGIRG}_{W,L}(1)$  happens whp. By (2.21), jointly, fixing the vertices  $v_n^1, v_n^2$  and letting  $K \rightarrow \infty$ ,

$$(T_K(v_n^1), T_K(v_n^2)) \xrightarrow{a.s.} (Y_1^E(v_n^1), Y_1^E(v_n^2)), \quad (2.32)$$

the joint distribution of the explosion times of  $v_n^1, v_n^2$  in  $\text{EGIRG}_{W,L}(1)$ . Observe that  $(T_K(v_n^1), T_K(v_n^2))$  is monotone increasing in  $K$  that will be relevant later on.

Next, we show in Proposition 3.8 the whp existence of a path connecting  $\tilde{v}_n^1(K), \tilde{v}_n^2(K)$  within  $\tilde{\mathcal{X}}_d(n)$  that has  $L$ -distance at most  $\varepsilon_K$  within  $\text{BGIRG}_{W,L}(n)$ . We do this using a subgraph created by *weight-dependent percolation*. That is, we keep edges  $e = (u, v) \in \mathcal{E}_B(n)$  if and only if their edge-length  $L_e \leq \text{thr}(W_u^{(n)}, W_v^{(n)})$ , for a function  $\text{thr}$  of  $W_u^{(n)}, W_v^{(n)}$ , that satisfies  $\text{thr}(x, y) = F_L^{(-1)}(p(x, y))$ , where we set  $p(x, y) \geq \exp\{-(\log x)^\gamma - (\log y)^\gamma\}$ ,  $\gamma \in (0, 1)$ . It follows that we keep an edge  $e = \{u, v\}$  with probability at least  $\exp\{-(\log W_u^{(n)})^\gamma - (\log W_v^{(n)})^\gamma\}$ . Let us denote the remaining subgraph by  $G^p$ . The advantage of choosing the percolation function like this is, as we show in Claim 6.2, that  $G^p$ , in distribution, is again a BGIRG with a new weight distribution  $W_p^{(n)}$  and new edge probabilities  $g_n^p$ , that again satisfy Assumption 2.2, 2.4 and 2.5, and the new parameters  $\alpha^p, \tau^p$  are unchanged, i.e.  $\alpha^p = \alpha, \tau^p = \tau$ .

Using this distributional identity, we describe a (doubly exponentially growing) boxing structure around  $\tilde{v}_n^1(K), \tilde{v}_n^2(K)$  in Lemma 6.3 that uses vertices and edges within  $G^p$ . Using this boxing structure, we construct a connecting path and control the weights of vertices along the connecting path between  $\tilde{v}_n^1(K), \tilde{v}_n^2(K)$ , as shown in Figures 3 and 4.

Since the percolation provides a deterministic upper bound on the edge-lengths of retained edges, once we have bounds on the weights of the vertices, we have a deterministic bound  $\varepsilon_K$  on the length of the connecting path, where  $\varepsilon_K$  tends to zero with  $K$ . The first part of the path, constructed by following the explosive path in  $\text{EGIRG}_{W,L}(1)$  is part of  $\text{BGIRG}_{W,L}(n)$  with probability that tends to 1 as  $n \rightarrow \infty$ , while the connecting path using the percolation and the boxing method exists with probability that tends to 1 as  $K \rightarrow \infty$ , as in Figure 4. Thus we arrive at the bound

$$d_L(v_n^1, v_n^2) \leq T_K(v_n^1) + T_K(v_n^2) + \varepsilon_K \leq Y_1^E(v_n^1) + Y_1^E(v_n^2) + \varepsilon_K, \quad (2.33)$$

where we have used the monotonicity of the limit in (2.32). From here we finish the proof by controlling the error probabilities by first choosing  $K$  large enough, then  $n$  large enough and showing that

$$(Y_1^E(v_n^1), Y_1^E(v_n^2)) \xrightarrow{d} (Y^{(1)}, Y^{(2)}), \quad (2.34)$$

two i.i.d. copies of  $Y_1^E(0)$ , a part that is non-trivial itself, similar to the method in the lower bound.



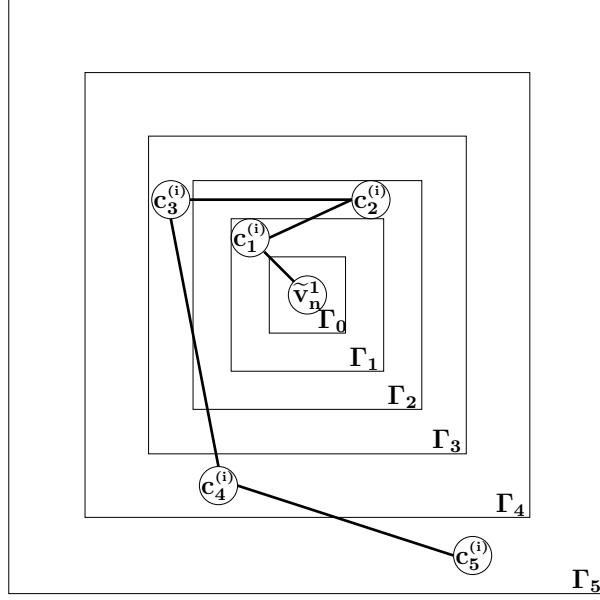


FIGURE 3. A schematic overview of the boxing structure around  $\tilde{v}_1^n$ . The radius of the annuli increases doubly-exponentially with their index. We drew a path connecting centres of subboxes within consecutive annuli, but we did not draw the subboxes. In the extended model, this construction can be continued indefinitely, yielding an infinite path. This is the idea of proving the existence of an infinite component.

**2.3. Structure of the paper.** In Section 3, we describe a coupling between  $\text{BGIRG}_{W,L}(n)$  and the extended model  $\text{EGIRG}_{W,L}(\lambda)$  and state two propositions capturing the upper and lower bound of Theorem 2.12. We provide the proof of Theorem 2.12 in this section subject to these propositions. In Section 4, we introduce a branching random walk in a random environment, and describe its behaviour when  $\mathbb{E}[W^2] < \infty$  or  $\mathbb{E}[W^2] = \infty$ . We also prove Theorem 2.11 in this section. The BRW allows us to bound the growth of neighbourhoods in  $\text{BGIRG}_{W,L}(n)$ , which we utilise in Section 5 by proving the lower bound in Theorem 2.12 (proof of Proposition 3.6 below). We discuss weight-dependent percolation and a boxing method in Section 6, both preliminaries for the upper bound in Theorem 2.12. Section 7 is devoted to the optimal explosive path in  $\text{EGIRG}_{W,L}(1)$  and its presence in  $\text{BGIRG}_{W,L}(n)$ , and contains the proof of the upper bound in Theorem 2.12 (proof of Proposition 3.8 below). We then use Section 8 to describe the necessary adjustments for our results to hold for scale-free percolation (proof of Theorem 2.14). Finally, in Section 9, we describe the connection between hyperbolic random graphs to GIRGs and show that Theorem 2.12 is valid for hyperbolic random graphs as well.

### 3. COUPLING OF GIRGS AND THEIR EXTENSION TO $\mathbb{R}^d$

In this section we state two main propositions for Theorem 2.12 and describe a *multiple-process coupling*. To be able to show the convergence of distances in (2.23) to the explosion time of  $\text{EGIRG}_{W,L}(1)$  (Definition 2.10), we need to relate  $\text{GIRG}_{W,L}(n)$  to  $\text{EGIRG}_{W,L}(1)$ . Therefore, we first blew the model up to  $\tilde{\mathcal{X}}_d(n)$ , obtaining a unit vertex-density model  $\text{BGIRG}_{W,L}(n)$ , and then extended this model to  $\mathbb{R}^d$ , obtaining  $\text{EGIRG}_{W,L}(\lambda)$ , using PPPs with intensity  $\lambda$  as vertex sets, allowing for vertex-densities other than 1. While in  $\text{BGIRG}_{W,L}(n)$  there are  $n$  vertices, the number of vertices in  $\text{EGIRG}_{W,L}(\lambda) \cap \tilde{\mathcal{X}}_d(n)$  is  $\text{Poi}(\lambda n)$ . Thus the two models cannot be compared directly. To circumvent this issue, we slightly increase/decrease the vertex-densities and even the edge-densities to obtain models with resp. more/less vertices in  $\tilde{\mathcal{X}}_d(n) = [-n^{1/d}/2, n^{1/d}/2]$  than  $n$ . Recall Definitions 2.8, 2.7.

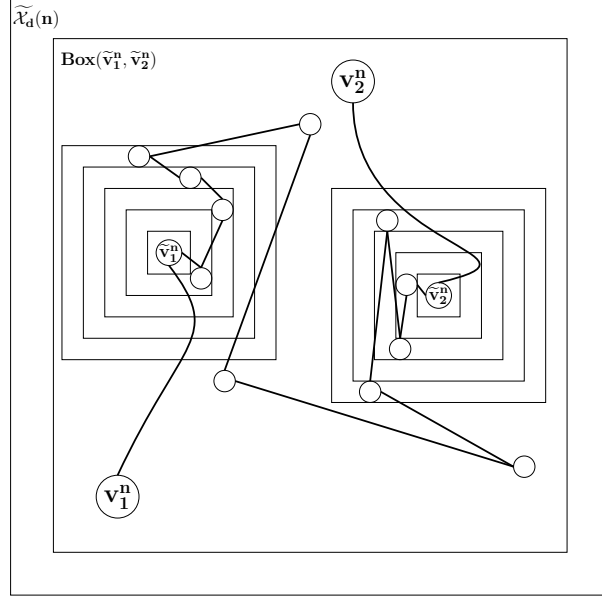


FIGURE 4. The idea of the proof of the upper bound. We connect  $v_n^1$  with  $v_n^2$  by following the two shortest explosive rays to two vertices with degree at least  $K$ , denoted by  $\tilde{v}_n^1(K), \tilde{v}_n^2(K)$ . Then we connect these two vertices using a boxing structure, and following two paths via centres of subboxes. We control the total weight on the edges of the connecting segment between  $\tilde{v}_n^1(K), \tilde{v}_n^2(K)$  by doing the boxing structure in a percolated graph, where only edges with small edge-weight are kept.

**Definition 3.1** (Extended GIRG with increased edge-density). *Let us define  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  as follows: Use the same vertex positions and vertex-weights  $(x, W_x)_{x \in \mathcal{V}_\lambda}$  as in the  $\text{EGIRG}_{W,L}(\lambda)$  model. Conditioned on  $(x, W_x)_{x \in \mathcal{V}_\lambda}$ , an edge between two vertices  $y, z \in \mathcal{V}_\lambda$  is present independently with probability*

$$\mathbb{P}(y \leftrightarrow z \text{ in } \overline{\text{EGIRG}}_{W,L}(\lambda) \mid (x, W_x)_{x \in \mathcal{V}_\lambda}) = \overline{C}_1 \overline{g}(y - z, W_y, W_z). \quad (3.1)$$

*Each edge  $e$  carries a random length  $L_e$ , i.i.d. from some distribution  $F_L$ . We denote the edge set by  $\overline{\mathcal{E}}_\lambda$ .*

An elementary claim is the following:

**Claim 3.2.** *Let  $V_\lambda^n \stackrel{d}{=} \text{Poi}(\lambda n)$  be the number of vertices in  $\mathcal{V}_\lambda \cap \tilde{\mathcal{X}}_d(n)$ . Set  $\xi_n := \sqrt{4 \log n / n}$ . Then there exists an  $n_0 \in \mathbb{N}$ , such that for all  $n \geq n_0$ ,  $V_{1-\xi_n}^n \leq n \leq V_{1+\xi_n}^n$  holds almost surely.*

*Proof.* The distribution of  $V_\lambda^n$  is  $\text{Poi}(n\lambda)$ . Let  $I_{1+\xi_n}(1) := (1 + \xi_n) - 1 - \log(1 + \xi_n) \leq \xi_n^2/2$  be the rate function of a Poisson random variable with parameter  $1 + \xi_n$  at 1. Then, using Chernoff bounds [47, Theorem 2.19],

$$\mathbb{P}(V_{1+\xi_n}^n < n) = \exp\{-n I_{1+\xi_n}(1)\} \leq \exp\{-n \xi_n^2/2\} = n^{-2}, \quad (3.2)$$

which is summable in  $n$ . Following similar steps for the  $1 - \xi_n$  case, the claim follows using the Borel-Cantelli lemma on the sequence of events  $V_{1-\xi_n}^n \leq n \leq V_{1+\xi_n}^n$ .  $\square$

**Claim 3.3.** *There exists a coupling between  $(\text{EGIRG}_{W,L}(1 \pm \xi_n))_{n \geq 3}$ ,  $\text{EGIRG}_{W,L}(1)$ , and the model  $(\text{BGIRG}_{W,L}(n))_{n \geq n_0}$ , where  $n_0$  is from Claim 3.2, such that for all  $n \geq 3$ ,*

$$\mathcal{V}_{1-\xi_n} \subseteq \mathcal{V}_1 \subseteq \mathcal{V}_{1+\xi_n}, \quad \mathcal{E}_{1-\xi_n} \subseteq \mathcal{E}_1 \subseteq \mathcal{E}_{1+\xi_n}, \quad (3.3)$$

and, for all  $n \geq n_0$ ,

$$(\mathcal{V}_{1-\xi_n} \cap \tilde{\mathcal{X}}_d(n)) \subseteq \mathcal{V}_B(n) \subseteq (\mathcal{V}_{1+\xi_n} \cap \tilde{\mathcal{X}}_d(n)). \quad (3.4)$$

*Proof.* We introduce a coupling between the PPPs  $(\mathcal{V}_{1\pm\xi_n})_{n\geq 3}$  for all  $n$  at once. Note that  $\xi_n$  is decreasing in  $n$  whenever  $n \geq 3$ . We use a PPP with intensity  $1 + \xi_3$  to create the vertex set  $\mathcal{V}_{1+\xi_3}$ . Then, for each vertex  $v \in \mathcal{V}_{1+\xi_3}$  we draw an i.i.d. uniform  $[0, 1]$  random variable that we denote by  $U_v$ . Finally, for all  $n \geq 3$  we set  $v \in \mathcal{V}_{1+\xi_n}$  if and only if  $U_v \leq (1 + \xi_n)/(1 + \xi_3)$ , and  $v \in \mathcal{V}_{1-\xi_n}$  if and only if  $U_v \leq (1 - \xi_n)/(1 + \xi_3)$ . Setting  $\xi_n = 0$  yields also a coupling to  $\text{EGIRG}_{W,L}(1)$ . The independent thinning of Poisson processes ensures that  $\mathcal{V}_{1\pm\xi_n}$  is a PPP with intensity  $1 \pm \xi_n$ . To determine the edge set of  $\text{EGIRG}_{W,L}(\lambda)$ , note that the edge probabilities are determined by the same function  $h$  for all  $\lambda$ . So, first we determine the presence of each possible edge between any two vertices in  $\mathcal{V}_{1+\xi_3}$ . This edge is present in  $\text{EGIRG}_{W,L}(1 \pm \xi_n)$  precisely when both end-vertices of the edge are present in  $\mathcal{V}_{1\pm\xi_n}$ . This coupling guarantees that the containment in (3.3) holds also for the edge-sets.

To construct a coupling ensuring (3.4), we determine  $\mathcal{V}_B(n)$ , given that  $V_{1+\xi_n}^n = k_1 \geq n$  and  $V_{1-\xi_n}^n = k_2 \leq n$ . Note that  $\mathcal{V}_B(n)$  is a set of  $n$  i.i.d. points sampled from  $\tilde{\mathcal{X}}_d(n)$ , while,  $\mathcal{V}_{1\pm\xi_n} \cap \tilde{\mathcal{X}}_d(n)$  conditioned on having  $k_1$  (resp.  $k_2$ ) points, has the distribution of a set of  $k_1$  (resp.  $k_2$ ) many i.i.d. points sampled from  $\tilde{\mathcal{X}}_d(n)$ . Thus, the way to generate the location of the  $n$  vertices in  $\text{BGIRG}_{W,L}(n)$  so that (3.3) holds is by taking the vertex set of  $\mathcal{V}_{1-\xi_n} \cap \tilde{\mathcal{X}}_d(n)$ , and then adding a uniform subset of points of size  $n - k_2$  from  $(\mathcal{V}_{1+\xi_n} \cap \tilde{\mathcal{X}}_d(n)) \setminus (\mathcal{V}_{1-\xi_n} \cap \tilde{\mathcal{X}}_d(n))$ .  $\square$

Now we couple the edge-sets of  $\text{BGIRG}_{W,L}(n)$  to the extended models. The content of the following claim is the precise version of the inequality in (2.27).

**Claim 3.4.** *There is a coupling of  $\text{BGIRG}_{W,L}(n)$  and  $\text{EGIRG}_{W,L}(\lambda), \overline{\text{EGIRG}}_{W,L}(\lambda)$  such that*

$$\mathcal{E}_B(n) \subseteq \bar{\mathcal{E}}_{1+\xi_n}, \quad \mathcal{E}_\lambda \subseteq \bar{\mathcal{E}}_\lambda, \quad (3.5)$$

where the latter holds for all  $\lambda > 0$ . Let  $E = (e_1, \dots, e_k) \subseteq \bar{\mathcal{E}}_{1+\xi_n}$  be a set of edges, where  $e_i$  connects vertices  $v_{i_1}, v_{i_2} \in \mathcal{V}_B(n)$  with locations and weights  $(x_{i_1}, W_{i_1}), (x_{i_2}, W_{i_2})$  in  $\text{EGIRG}_{W,L}(1 + \xi_n)$ . Assume that the event  $E_{\Delta,W} := \{\forall i \leq k : \|x_{i_1} - x_{i_2}\| \in I_\Delta(n), W_{i_1}, W_{i_2} \in I_w(n)\}$  holds and let  $n \geq n_0$  where  $n_0$  is from Claim 3.2. Then,

$$\mathbb{P}\left((E \cap \mathcal{E}_B(n)) \ominus (E \cap \mathcal{E}_{1+\xi_n}) \neq \emptyset \mid E \subseteq \bar{\mathcal{E}}_{1+\xi_n}, E_{\Delta,W}\right) \leq k(\epsilon(n) + 2\epsilon_{\text{TV}}(n)) \quad (3.6)$$

where  $\ominus$  denotes symmetric difference. The same bound is true for  $\text{EGIRG}_{W,L}(1 - \xi_n)$  when we assume that  $v_{i_1}, v_{i_2} \in \mathcal{V}_{1-\xi_n}$ .

*Proof.* Recall the notations  $g_n^B, h, \bar{g}$  from (2.15), from in Assumption 2.5 and from Claim 2.6, the bound from (2.16), and recall the weights  $W^{(n)}, W$  from Assumption 2.4 and their total variation distance from (2.9). By the coupling described in Claim 3.3, the vertex set of  $\text{BGIRG}_{W,L}(n)$  is a subset of  $\mathcal{V}_{1+\xi_n}$ . For any possible edge connecting vertices in  $\mathcal{V}_B(n)$ , we construct a coupling between the presence of this edge in  $\text{BGIRG}_{W,L}(n)$  vs that of in  $\mathcal{E}_{1+\xi_n}, \bar{\mathcal{E}}_{1+\xi_n}$  and  $\mathcal{E}_{1-\xi_n}$  (in the latter case, given that the two vertices are also part of  $\mathcal{V}_{1-\xi_n}$ ). First, use the optimal coupling realising the total-variation distance between  $W^{(n)}$  and  $W$  for each endpoint of the edges under consideration. Then, for an edge  $e$  connecting vertices  $i_1, i_2$ ,

$$\mathbb{P}(W_{i_1}^{(n)} \neq W_{i_1} \text{ or } W_{i_2}^{(n)} \neq W_{i_2}) = 2\epsilon_{\text{TV}}(n). \quad (3.7)$$

Given that the two pairs of weights are equal, for a possible edge  $e$  connecting vertices with locations and weights  $(x_{i_1}, W_{i_1}), (x_{i_2}, W_{i_2})$ , draw  $U_e$ , an i.i.d. rv uniform in  $[0, 1]$ . Include this edge respectively in  $\mathcal{E}_B(n), \mathcal{E}_{1\pm\xi_n}, \bar{\mathcal{E}}_{1+\xi_n}$ , if and only if

$$\begin{aligned} e \in \mathcal{E}_B(n) &\Leftrightarrow U_e \leq g_n^B(x_{i_1}, x_{i_2}, (W_j)_{j \in [n] \setminus \{i_1, i_2\}}) =: g_n^B(e), \\ e \in \mathcal{E}_{1\pm\xi_n}(n) &\Leftrightarrow U_e \leq h(x_{i_1} - x_{i_2}, W_{i_1}, W_{i_2}) =: h(e) \\ e \in \bar{\mathcal{E}}_{1+\xi_n}(n) &\Leftrightarrow U_e \leq \bar{C}_1 \bar{g}(x_{i_1} - x_{i_2}, W_{i_1}, W_{i_2}) =: \bar{C}_1 \bar{g}(e) \end{aligned} \quad (3.8)$$

These, per definitions of the models in Def. 2.8, 2.7, 3.1 give the right edge-probabilities. By (2.16) and Claim 2.6, both  $h(e), g^B(e) \leq \bar{C}_1 \bar{g}(e)$ . Since  $\mathcal{V}_B(n) \subset \mathcal{V}_{1+\xi_n}$  and the vertex set is  $\mathcal{V}_\lambda$  for both

$\text{EGIRG}_{W,L}(\lambda), \overline{\text{EGIRG}}_{W,L}(\lambda)$ , this, together with (3.8) ensures (3.5). By Assumption 2.5, for any  $e$ ,

$$\begin{aligned} \mathbb{P}((e \in \mathcal{E}_B(n)) \ominus (e \in \mathcal{E}_{1+\xi_n}) \neq \emptyset \mid e \in \bar{\mathcal{E}}_{1+\xi_n}, E_{\Delta,W}) &\leq 2\epsilon_{\text{TV}}(n) + \frac{|g_n^B(e) - h(e)|}{C_1 \bar{g}(e)} \\ &\leq 2\epsilon_{\text{TV}}(n) + \epsilon(n), \end{aligned} \quad (3.9)$$

by Assumption 2.5 and Claim 2.6. A union bound finishes the proof of (3.6).  $\square$

*Key propositions for Theorem 2.12.* Next, we formulate three propositions. The first shows the existence of an infinite component in  $\text{EGIRG}_{W,L}(\lambda)$ , and the other two formulate the upper and lower bound on  $d_L(v_n^1, v_n^2)$  described in the overview in (2.30) and (2.33). Recall the model-dependent notation for metric balls from Definition 2.9.

**Proposition 3.5** (Existence of  $\mathcal{C}_\infty^\lambda$ ). *Consider the  $\text{EGIRG}_{W,L}(\lambda)$  model, as in Definition 2.7, with power-law parameter  $\tau \in (2, 3)$ . Then, for all  $\lambda > 0$ , there is a unique infinite component  $\mathcal{C}_\infty^\lambda$ .*

We mention that there are some existing results also for  $\tau \geq 3$ . In this case, for a specific (illustrative) choice of the edge probabilities, the infinite model  $\text{EGIRG}_{W,L}(\lambda)$  has an infinite component only when the edge-intensity (or equivalently, the vertex intensity) is sufficiently high, see [32, 34]. For one dimension, hyperbolic random graphs form a special case when there is never a giant component when  $\tau \geq 3$ , see [18].

Recall from before Theorem 2.12 that  $\mathcal{C}_{\max}$  denotes the unique, linear-sized giant component of  $\text{BGIRG}_{W,L}(n)$ , that exists whp by Bringmann *et al.* [22, Theorem 2.2].

**Proposition 3.6** (Lower bound on distances in the BGIRG model). *Let  $v_n^q$ ,  $q = 1, 2$ , be two uniformly chosen vertices in  $\mathcal{C}_{\max}$  of  $\text{BGIRG}_{W,L}(n)$ , with parameters  $\tau \in (2, 3)$ ,  $\alpha > 1$  and  $d \geq 1$ . For  $q \in \{1, 2\}$ , let*

$$V_+^q(n, k) := d_L^{1+\xi_n}(v_n^q, \partial B_{1+\xi_n}^G(v_n^q, k)), \quad (3.10)$$

*be defined in  $\text{EGIRG}_{W,L}(1 + \xi_n)$ . Let  $A_k$  be the event in (2.29). Then, there exists a sequence  $k_n$  with  $k_n \rightarrow \infty$ , such that  $\mathbb{P}(A_{k_n}) \rightarrow 1$  as  $n \rightarrow \infty$ , and*

$$d_L(v_n^1, v_n^2) \geq \mathbb{1}_{A_{k_n}} (V_+^1(n, k_n) + V_+^2(n, k_n)), \quad (3.11)$$

*and finally, conditionally on  $A_{k_n}$ ,  $V_+^1(n, k_n), V_+^2(n, k_n)$  are independent.*

In the same manner, we formulate a proposition for the upper bound. Here, we shall take a double limit approach, so the error bounds are somewhat more involved. For a vertex  $v \in \mathcal{V}_1$ , let  $\tilde{v}(K)$  be the first vertex with weight larger than  $K$  on the optimal explosive path  $\pi_{\text{opt}}^1(v)$  from (2.20) in  $\text{EGIRG}_{W,L}(1)$ , let  $\pi_{\text{opt}}^1[u, \tilde{u}(K)]$  denote the segment of  $\pi_{\text{opt}}^1(u)$  from  $u$  to  $\tilde{u}(K)$ , and let  $T_K(u) := d_L^1(u, \tilde{u}(K)) = |\pi_{\text{opt}}^1[u, \tilde{u}(K)]|_L$ .

**Lemma 3.7** (Best explosive path can be followed). *Let  $v_n^1, v_n^2$  be uniformly chosen vertices in  $\mathcal{C}_{\max}$  of  $\text{BGIRG}_{W,L}(n)$ . Set*

$$\tilde{A}_{n,K} := \{\pi_{\text{opt}}^1[v_n^q, \tilde{v}_n^q(K)] \text{ is present in } \text{BGIRG}_{W,L}(n), \text{ for } q \in \{1, 2\}\}. \quad (3.12)$$

*Then  $1 - \mathbb{P}(\tilde{A}_{n,K}) \leq f(n, K)$  for some function  $f(n, K)$  with  $\lim_{n \rightarrow \infty} f(n, K) = 0$  for each fixed  $K$ .*

Note that  $\tilde{A}_{n,K}$  implies the event that  $v_n^1, v_n^2$  are in  $\mathcal{C}_\infty^1$  of  $\text{EGIRG}_{W,L}(1)$ , otherwise  $\pi_{\text{opt}}^1(v_n^q)$  is not defined. With the event  $\tilde{A}_{n,K}$  at hand, we are ready to state the key proposition for the upper bound.

**Proposition 3.8** (Upper bound on distances in the GIRG model in the explosive case). *Let  $v_n^1, v_n^2$  be uniformly chosen vertices in  $\mathcal{C}_{\max}$  of  $\text{BGIRG}_{W,L}(n)$ . On  $\tilde{A}_{n,K}$  from (3.12), for some  $\epsilon_K$  that tends to 0 as  $K \rightarrow \infty$ , the bound*

$$d_L(v_n^1, v_n^2) \leq T_K(v_n^1) + T_K(v_n^2) + \epsilon_K \quad (3.13)$$

*holds with probability  $1 - \eta(K)$ , for some function  $\eta(K) \rightarrow 0$  as  $K \rightarrow \infty$ .*

We note that the short connection between the optimal explosive paths is the key part that is missing from [51, Conjecture 1.11]. We prove Proposition 3.5 in Section 6, Proposition 3.6 in Section 5 and Lemma 3.7 and Proposition 3.8 in Section 7. From Lemma 3.7, Propositions 3.6 and 3.8, the proof of Theorem 2.12 follows

*Proof of Theorem 2.12 subject to Propositions 3.6 and 3.8.* Let  $x$  be a continuity point of the distribution function of  $Y^{(1)} + Y^{(2)}$ . We start by bounding  $\mathbb{P}(d_L(v_n^1, v_n^2) \leq x)$  from above by using Proposition 3.6. Let  $V_1^q(n, k) := d_L^1(v_n^q, \partial B_1^G(v_n^q, k))$ , similar to (3.10). Due to the coupling developed in Claim 3.3, given that  $v_n^q \in \mathcal{V}_1$ ,  $\mathcal{E}_1 \subseteq \mathcal{E}_{1+\xi_n}$  holds. Thus  $V_+^q(n, k_n) \leq V_1^q(n, k_n)$ . Let  $\pi_\star^q(k)$  be the shortest path connecting  $v_n^q$  to  $\partial B_{1+\xi_n}^G(v_n^q, k)$ . Since on  $A_{k_n}$ ,  $\pi_\star^q(k_n) \subseteq B_1^G(v_n^q, k_n)$ , this means that conditioned on  $A_{k_n}$ ,

$$(V_1^1(n, k_n), V_1^2(n, k_n)) = (V_+^1(n, k_n), V_+^2(n, k_n)) \quad (3.14)$$

Further,  $\pi_\star^q(k) \subseteq \text{BGIRG}_{W,L}(n)$  and  $B_n^G(v_n^1, k) \cap B_n^G(v_n^2, k) = \emptyset$  implies that the shortest path between  $v_n^1, v_n^2$  in  $\text{BGIRG}_{W,L}(n)$  is at least as long as the total length of  $\pi_\star^1(k)$  and  $\pi_\star^2(k)$ . Thus, conditioned on  $A_{k_n}$ ,

$$\begin{aligned} \mathbb{P}(d_L(v_n^1, v_n^2) \leq x) &\leq \mathbb{P}(\mathbb{1}_{A_{k_n}} (V_+^1(n, k_n) + V_+^2(n, k_n)) \leq x) \\ &= \mathbb{P}(\mathbb{1}_{A_{k_n}} (V_1^1(n, k_n) + V_1^2(n, k_n)) \leq x) \\ &\leq 1 - \mathbb{P}(A_{k_n}^c) + \mathbb{P}(V_1^1(n, k_n) + V_1^2(n, k_n) \leq x), \end{aligned} \quad (3.15)$$

where  $1 - \mathbb{P}(A_{k_n}^c) \rightarrow 0$  by Proposition 3.6. Next we show that the joint distribution of these two variables tend to  $(Y^{(1)}, Y^{(2)})$  as  $n \rightarrow \infty$ . The model  $\text{EGIRG}_{W,L}(1)$  is translation invariant, thus, marginally,

$$V_q^1(n, k_n) \stackrel{d}{=} d_L^1(0, \partial B_1^G(0, k_n)) =: X_{k_n}, \quad (3.16)$$

for  $q \in \{1, 2\}$ , where we condition on  $0 \in \mathcal{V}_1$ . By (3.14) and the independence of  $V_1^+(n, k_n)$  and  $V_2^+(n, k_n)$  that both hold on  $A_{k_n}$ ,  $V_1^1(n, k_n)$  and  $V_2^1(n, k_n)$  are also (conditionally) independent of each other. Thus, the distributional identity

$$(V_1^1(n, k_n), V_1^2(n, k_n)) \stackrel{d}{=} (X_{k_n}^{(1)}, X_{k_n}^{(2)}) \quad (3.17)$$

holds on  $A_{k_n}$ , where  $X_{k_n}^{(q)}$  are i.i.d. copies of  $X_{k_n}$ . Finally, we show that

$$X_{k_n} = d_L^1(0, \partial B_1^G(0, k_n)) \xrightarrow{a.s.} Y_1^E(0), \quad (3.18)$$

the (possibly infinite) explosion time<sup>2</sup> of 0 in  $\text{EGIRG}_{W,L}(1)$ . This is actually a general statement that holds in any locally finite graph: the weighted distance to the boundary of the graph-distance ball of radius  $k$  tends to the explosion time, as the radius tends to infinity.

Recall that  $Y_1^E(0) = \lim_{k \rightarrow \infty} \tau_1^E(0, k)$  from Definition 2.10, the distance to the  $k$ -th closest vertex. Let  $\tau(k) := \tau_1^E(0, k)$ . To show that  $Y_1^E(0) = \lim_{k \rightarrow \infty} d_L^1(0, \partial B_1^G(0, k))$ , we show that for any  $t > 0$ , the events  $\{Y_1^E(0) \leq t\} = \{\lim_{k \rightarrow \infty} \tau(k) \leq t\}$  and  $\{\lim_{k \rightarrow \infty} d_L^1(0, \partial B_1^G(0, k)) \leq t\} = \{\lim_{k \rightarrow \infty} X_k \leq t\}$  coincide. To show this, we first observe that

$$\{\lim_{k \rightarrow \infty} \tau(k) \leq t\} = \{|B_1^L(0, t)| = \infty\}.$$

Indeed, when  $|B_1^L(0, t)| = \infty$ , then  $\tau(k)$ , the  $L$ -distance of  $k$ th closest vertex to 0 is also closer than  $t$ , hence  $\tau(k) \leq t$  holds. Reversely, assuming  $|B_1^L(0, t)| =: K < \infty$  implies that  $\tau(K+1) > t$ .

Next we show that  $\{|B_1^L(0, t)| = \infty\} \subseteq \{\lim_{k \rightarrow \infty} d_L^1(0, \partial B_1^G(0, k)) \leq t\}$ . Since all the degrees are finite, if there are infinitely many vertices within  $L$ -distance  $t$ , then there must be also a vertex  $v$  at graph distance  $k$  that is within  $L$ -distance  $t$ , (simply since one cannot squeeze in infinitely many vertices within graph distance  $k-1$ ). Hence, the distance between 0 and  $\partial B_1^{L,G}(0, k)$  is at most  $t$ . This holds for all  $k \geq 0$ , hence,

$$\{|B_1^L(0, t)| = \infty\} \subseteq \bigcap_{k \in \mathbb{N}} \{\exists v \in \partial B_1^G(0, k), d_L(0, v) \leq t\} = \bigcap_{k \in \mathbb{N}} \{X_k \leq t\} = \{\lim_{k \rightarrow \infty} X_k \leq t\}, \quad (3.19)$$

<sup>2</sup>The definition  $\lim_{k \rightarrow \infty} X_k$  is not the standard definition of the explosion time given in Definition 2.10 so we have to show that the two limits are the same.

where the last equality holds since  $X_k$  is non-decreasing in  $k$ . For the reverse direction, observe that  $|B_1^L(0, t)| = \sum_{k=1}^{\infty} |\partial B_1^G(0, k) \cap B_1^L(0, t)|$ , and if  $|B_1^L(0, t)| =: Z < \infty$ , then only finitely many terms can be non-zero. If a term with index  $k$  is non-zero then so is every smaller index  $i \leq k$ , since, the term being non-zero implies that there is a vertex  $v$  in  $\partial B_1^G(0, k)$  with  $d_L(0, v) \leq t$ . But then any vertex  $w$  on the shortest path from  $v$  to 0 has  $d_L(0, w) \leq t$ . For any term  $k$  that is 0,  $|\partial B_1^G(0, k) \cap B_1^L(0, t)| = 0$  implies that every vertex at graph distance  $k$  is at weighted distance larger than  $t$ , hence  $d_L^1(0, \partial B_1^G(0, k)) = X_k > t$ . Hence the limit of  $X_k$  is also  $> t$  in this case. This finishes the argument that  $\lim_{k \rightarrow \infty} X_{k_n} = Y_1^E(v)$ .

Returning to (3.15), combined with (3.17) and this convergence result finishes the proof that  $\mathbb{P}(d_L(v_n^1, v_n^2) \leq x) \leq \mathbb{P}(Y_1^{(1)} + Y_1^{(2)} \leq x) + \varepsilon$  for all  $\varepsilon > 0$  and sufficiently large  $n$ , where  $Y_1^{(1)}, Y_1^{(2)}$  are i.i.d. copies of  $Y_1^E(0)$ .

To estimate  $\mathbb{P}(d_L(v_n^1, v_n^2) \leq x)$  from below, we use Proposition 3.8. Recall the event  $\tilde{A}_{n,K}$  from (3.12). Let us write

$$A_{n,K}^* := \tilde{A}_{n,K} \cap \{d_L(v_n^1, v_n^2) \leq T_K(v_n^1) + T_K(v_n^2) + \varepsilon_K\}.$$

By the statement of Proposition 3.8 and Lemma 3.7,  $\mathbb{P}(A_{n,K}^*) \geq 1 - f(n, K) - \eta(K)$ . Further, since  $T_K(v_n^q)$  denotes the  $L$ -length of a section on the optimal path to infinity, (see the definition before Lemma 3.7)  $T_K(v_n^q) \leq Y_1^E(v_n^q)$ , the explosion time of  $v_n^q$  in  $\text{EGIRG}_{W,L}(1)$ . Thus

$$\mathbb{P}(d_L(v_n^1, v_n^2) \leq x) \geq (1 - f(n, K) - \eta(K))\mathbb{P}(Y_1^E(v_n^1) + Y_1^E(v_n^2) + \varepsilon_K \leq x). \quad (3.20)$$

We will show below that for two i.i.d. copies  $Y^{(1)}, Y^{(2)}$  of the explosion time of 0 in  $\text{EGIRG}_{W,L}(1)$ ,

$$(Y_1^E(v_n^1), Y_1^E(v_n^2)) \xrightarrow{d} (Y^{(1)}, Y^{(2)}). \quad (3.21)$$

Given this, fix  $\delta, x$  first and then choose  $K_1$  such that for all  $K \geq K_1$ ,  $\eta(K) \leq \delta/4$ . Let us assume that  $x$  is a continuity point of the distribution of  $Y^{(1)} + Y^{(2)}$ , and  $K_2 \geq K_1$  satisfies that for all  $K \geq K_2$ ,

$$|\mathbb{P}(Y^{(1)} + Y^{(2)} \leq x - \varepsilon_K) - \mathbb{P}(Y^{(1)} + Y^{(2)} \leq x)| \leq \delta/4. \quad (3.22)$$

If a function is continuous at  $x$ , then it is also continuous on some interval  $(x - \zeta, x]$ , for some small  $\zeta > 0$ , so let  $K_3 \geq K_2$  satisfy further that  $\varepsilon_{K_3} \leq \zeta$ . Thus,  $x - \varepsilon_{K_3}$  is also a continuity point of the distribution of  $Y^{(1)} + Y^{(2)}$ . Let  $n_1$  now satisfy that for all  $n \geq n_1$ ,

$$|\mathbb{P}(Y_1^E(v_n^1) + Y_1^E(v_n^2) \leq x - \varepsilon_{K_3}) - \mathbb{P}(Y^{(1)} + Y^{(2)} \leq x - \varepsilon_{K_3})| \leq \delta/4. \quad (3.23)$$

Note that  $f(n, K)$  for fixed  $K$  tends to zero with  $n$  (see Lemma 3.7). So, set  $n_2 := n_2(K_3) \geq n_1$  such that for all  $n \geq n_2$ ,  $f(n, K_3) \leq \delta/4$ . Combining (3.20)-(3.23) yields that for all  $n \geq n_2$ ,

$$\begin{aligned} \mathbb{P}(d_L(v_n^1, v_n^2) \leq x) &\geq (1 - f(n_2, K_3) - \eta(K_3)) \left( \mathbb{P}(Y^{(1)} + Y^{(2)} \leq x - \varepsilon_{K_3}) \pm \delta/2 \right) \\ &\geq \mathbb{P}(Y^{(1)} + Y^{(2)} \leq x) - \delta. \end{aligned} \quad (3.24)$$

This finishes the proof of the lower bound, given (3.21). Next we show (3.21). Recall from the proof of the upper bound that on the event  $A_{k_n}$ , as in (2.29), the variables  $V_q^1(n, k_n) = d_L^1(v_n^q, \partial B_1^{G,L}(v_n^q, k))$  are independent variables for  $q \in \{1, 2\}$ , and they satisfy the distributional identity (3.17) on  $A_{k_n}$  and they approximate, as  $k_n \rightarrow \infty$ , two i.i.d. copies of  $Y_1^E(0)$ . Let both  $z_1$  and  $z_2$  be continuity points of the distribution of  $Y_1^E(0)$ . Thus, estimating the joint distribution functions of (3.21) by the triangle inequality, we write

$$\begin{aligned} &|\mathbb{P}(Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2) - \mathbb{P}(Y^{(1)} \leq z_1, Y^{(2)} \leq z_2)| \\ &\leq |\mathbb{P}(Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2) - \mathbb{P}(V_1^1(n, k_n) \leq z_1, V_2^1(n, k_n) \leq z_2)| \\ &\quad + |\mathbb{P}(V_1^1(n, k_n) \leq z_1, V_2^1(n, k_n) \leq z_2) - \mathbb{P}(X_{k_n}^{(1)} \leq z_1, X_{k_n}^{(2)} \leq z_2)| \\ &\quad + |\mathbb{P}(X_{k_n}^{(1)} \leq z_1, X_{k_n}^{(2)} \leq z_2) - \mathbb{P}(Y^{(1)} \leq z_1, Y^{(2)} \leq z_2)| \\ &=: T_1 + T_2 + T_3, \end{aligned} \quad (3.25)$$



where we denote the terms on the rhs in the second, third and fourth row by  $T_1, T_2$  and  $T_3$ , respectively. Then, due to the distributional identity (3.17) on  $A_{k_n}$ ,  $T_2$  is bounded from above by  $1 - \mathbb{P}(A_{k_n})$ , and  $T_3$  tends to 0 by the argument after (3.19) as  $k_n \rightarrow \infty$  (that is,  $n \rightarrow \infty$ ). To estimate  $T_1$ , note that the four variables involved are all defined on the same probability space (namely, in terms of the realisation  $\text{EGIRG}_{W,L}(1)$ ). Further, fixing  $n$ , the variable  $V_q^1(n, k) = d_L^1(v_n^q, \partial B_1^{G,L}(v_n^q, k))$  is increasing in  $k$  and tends to  $Y_1^E(v_n^1)$ , the explosion time of  $v_n^q$  in a realisation of  $\text{EGIRG}_{W,L}(1)$  as  $k \rightarrow \infty$ . In particular, for every  $k \geq 0$ , almost surely,  $Y_1^E(v_n^1) - V_q^1(n, k) \geq 0$  holds. Hence

$$\begin{aligned} T_1 &= |\mathbb{P}(Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2) - \mathbb{P}(V_1^1(n, k_n) \leq z_1, V_2^1(n, k_n) \leq z_2)| \\ &= \mathbb{P}(V_1^1(n, k_n) \leq z_1, V_2^1(n, k_n) \leq z_2) - \mathbb{P}(Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2) \end{aligned} \quad (3.26)$$

Next we observe that  $Y_1^E(v_n^1) - V_q^1(n, k) \geq 0$  tends to zero as we increase  $k$ , so the difference is unlikely to be less than  $\delta > 0$  for  $k$  large enough. Thus, we can add for any  $\delta > 0$  the extra term in the middle:

$$\begin{aligned} T_1 &\leq \mathbb{P}(V_1^1(n, k_n) \leq z_1, V_2^1(n, k_n) \leq z_2) - \mathbb{P}(V_1^1(n, k_n) \leq z_1 - \delta, V_2^1(n, k_n) \leq z_2 - \delta) \\ &\quad + \mathbb{P}(V_1^1(n, k_n) \leq z_1 - \delta, V_2^1(n, k_n) \leq z_2 - \delta) - \mathbb{P}(Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2). \end{aligned} \quad (3.27)$$

Note that the first row can be bounded from above by the probability of the event that at least one of the variables  $V_q^1(n, k_n)$  falls in a small interval of length  $\delta$  while the second row equals the probability below:

$$\begin{aligned} T_1 &\leq \mathbb{P}(V_1^1(n, k_n) \in (z_1 - \delta, z_1]) + \mathbb{P}(V_1^1(n, k_n) \in (z_2 - \delta, z_2]) \\ &\quad + \mathbb{P}(\{V_1^1(n, k_n) \leq z_1 - \delta, V_2^1(n, k_n) \leq z_2 - \delta\} \setminus \{Y_1^E(v_n^1) \leq z_1, Y_1^E(v_n^2) \leq z_2\}) \end{aligned} \quad (3.28)$$

The event in the second row is satisfied only when  $Y_q^E(v_n^1) - V_q^1(n, k_n) \geq \delta$  for one of the  $q = 1, 2$ 's. Due to the translation invariance of the model, we can use that *marginally*  $(V_q^1(n, k_n), Y_1^E(v_n^q))$  has the same distribution as  $(X_{k_n}, Y_1^E(0))$  from (3.16), and hence, by a union bound, the second row in (3.28) is at most

$$2\mathbb{P}(Y_1^E(0) - X_{k_n} \geq \delta), \quad (3.29)$$

which tends to 0 for every fixed  $\delta > 0$  as  $k_n \rightarrow \infty$  by the argument after (3.17). To bound the first row on the rhs of (3.28), due to the translation invariance again, marginally,  $V_q^1(n, k_n) = X_{k_n}^{(q)} \stackrel{d}{=} X_{k_n}$  and  $X_{k_n} \stackrel{d}{\rightarrow} Y_1^E(0)$ , and  $z_1, z_2$  are continuity points of the distribution of  $Y_1^E(0)$ . Hence, for any  $\varepsilon > 0$ , one can choose first  $\delta > 0$  small enough and then  $k_n$  large enough so that this probability is less than  $\varepsilon$ .  $\square$

#### 4. BRANCHING RANDOM WALKS IN RANDOM ENVIRONMENT AND GIRGS

In this section, we provide the framework for the lower bound in Theorem 2.12. As a preparation for Proposition 3.8, we show that every path with finite total  $L$ -length to infinity must leave the subgraph of  $\text{EGIRG}_{W,L}(\lambda)$  restricted to vertices with bounded weight. Both proofs rely on a coupling to a process, a *branching random walk* (BRW), that has a faster growing neighbourhood around a vertex  $v$  than the infinite component of  $\text{EGIRG}_{W,L}(\lambda)$ . Recall  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  from Def. 3.1. This BRW has a random environment. The environment is formed by  $(x, W_x)_{x \in \mathcal{V}_\lambda}$ , where  $\mathcal{V}_\lambda$  is a PPP of intensity  $\lambda$  on  $\mathbb{R}^d$  and  $(W_x)_{x \in \mathcal{V}_\lambda}$  are i.i.d. copies of  $W$ . We name the individuals in the *skeleton* of the BRW (a random branching process tree without any spatial embedding), via the Harris-Ulam manner. That is, children of a vertex are sorted arbitrarily and we call the root  $\emptyset$ , its children  $1, 2, \dots$ , their children  $11, 12, \dots, 21, 22, \dots$  and so on. Generally, individual  $\underline{i} := i_1 i_2 \dots i_k$  is the  $i_k^{\text{th}}$  child of the  $i_{k-1}^{\text{th}}$  child  $\dots$  of the  $i_1^{\text{th}}$  child of the root. This coding for all individuals is referred to as the *name* of an individual. We write  $p(\underline{i})$  for parent of individual  $\underline{i}$ , set  $p(\emptyset) = \emptyset$ . Abusing notation somewhat, we write  $x_{\underline{i}}, W_{\underline{i}}$  for the location and weight of the individual  $\underline{i}$ .

We locate the root  $\emptyset$  at the initial vertex  $v \in \mathcal{V}_\lambda$ , thus we set  $x_\emptyset := v$ . Conditioned on the environment, let the number of children at location  $x \in \mathcal{V}_\lambda \setminus \{x_{\underline{i}}\}$  of every individual  $\underline{i}$  be *independent*, and denoted and distributed as

$$N_x^B(\underline{i}) \stackrel{d}{=} \text{Ber}(\overline{C}_1 \min\{1, \bar{a}_1 \|x_{\underline{i}} - x\|^{-\alpha d} (W_{\underline{i}} W_x)^\alpha\}) = \text{Ber}(\overline{C}_1 \bar{g}(x_{\underline{i}} - x, W_{\underline{i}}, W_x)). \quad (4.1)$$

Then, the location and weight of the offspring of  $\underline{i}$  in the Bernoulli BRW is described as

$$\mathcal{N}^B(\underline{i}) := \bigcup_{x \in \mathcal{V}_\lambda \setminus \{x_{\underline{i}}\}: N_x^B(\underline{i})=1} \{(x, W_x)\}. \quad (4.2)$$

Note that individuals at the same location reproduce conditionally independently. We denote the resulting BRW by  $\text{BerBRW}_\lambda(v)$  and its generation sets by  $\mathcal{G}_\lambda^{\text{Ber}}(v, k)$ ,  $k \in \mathbb{N}$  and  $v$  the root. Note that the Bernoulli BRW is created so that the probability that a particle at a location  $y \in \mathcal{V}_\lambda$  has a child at location  $z \in \mathcal{V}_\lambda \setminus \{y\}$  equals the probability of the presence of the edge  $(y, z)$  in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$ , namely

$$\begin{aligned} \mathbb{P}(\text{an individual at } y \text{ has a child at } z \text{ in } \text{BerBRW}_\lambda(v) \mid (x, W_x)_{x \in \mathcal{V}_\lambda}) \\ = \mathbb{P}((y, z) \in \overline{\mathcal{E}}_\lambda \mid (x, W_x)_{x \in \mathcal{V}_\lambda}). \end{aligned} \quad (4.3)$$

After  $\text{BerBRW}_\lambda(v)$  is generated, we assign i.i.d. edge-lengths from distribution  $F_L$  to all existing parent-child relationships  $(p(\underline{i}), \underline{i})$ . We denote the resulting edge-weighted BRW by  $\text{BerBRW}_{\lambda,L}(v)$ . Let us denote by  $\overline{B}_\lambda^L(v, t)$ ,  $\overline{B}_\lambda^G(v, t)$  the set of vertices within  $L$ -distance  $t$  and graph distance  $t$  of  $v$  in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$ , and  $B_{\text{Ber}}^L(v, t)$ ,  $B_{\text{Ber}}^G(v, t)$  the corresponding quantities in  $\text{BerBRW}_{\lambda,L}(v)$ , that is, the set of individuals available from  $\emptyset = v$  on paths of  $L$ -length at most  $t$  and the set of individuals with name-length (= generation number) at most  $t$ , respectively. The exploration process on a similar BRW (defined in terms of the SFP model), coupled to the exploration on the SFP model is introduced in [51, Section 4.2]. The exploration process on  $\text{BerBRW}_{\lambda,L}(v)$  (as defined above) coupled to the exploration on  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  is analogous. The exploration algorithm runs on  $\text{BerBRW}_{\lambda,L}(v)$  and  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  at the same time, exploring  $B_{\text{Ber}}^L(v, t)$  as  $t$  increases, and thinning parent-child relationships in  $\text{BerBRW}_{\lambda,L}(v)$  that are between already explored locations, yielding  $\overline{B}_\lambda^L(v, t)$ . A consequence of this thinning procedure is the following lemma, which is an adaptation and simplified statement of [51, Proposition 4.1]:

**Lemma 4.1.** *There is a coupling of the exploration of  $(B_{\text{Ber}}^L(v, t))_{t \geq 0}$  on  $\text{BerBRW}_{\lambda,L}(v)$  to the exploration of  $(\overline{B}_\lambda(v, t))_{t \geq 0}$  on  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  so that for all  $t \geq 0$ , under the coupling,*

$$\overline{B}_\lambda^L(v, t) \subseteq B_{\text{Ber}}^L(v, t). \quad (4.4)$$

Further, there is a coupling of  $\text{BerBRW}_\lambda(v)$  and  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  such that under the coupling,

$$\partial \overline{B}_\lambda^G(v, t) \subseteq \partial B_{\text{Ber}}^G(v, t), \quad (4.5)$$

i.e. the vertices at graph distance  $k$  from  $v$  in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  form a subset of the individuals in generation  $k$  of  $\text{BerBRW}_\lambda(v)$ .

*Proof.* The proof of Lemma 4.1 is a direct analogue of the proof of [51, Proposition 4.1] and we refer the interested reader to their proof. The only difference is that in their proof the vertex set is  $\mathbb{Z}^d$  while here it is a PPP  $\mathcal{V}_\lambda$ . Heuristically, the coupling follows by (4.3), but it is non-trivial since one needs to assign the edge-lengths to parent-child relationships in  $\text{BerBRW}_\lambda(v)$  as well as the edge-lengths  $L_e$  in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  also in a coupled way.  $\square$

A combination of the couplings in Claim 3.4 and Lemma 4.1 yields that for all  $t$ , under the couplings, for all  $\lambda \geq 1 + \xi_n$ ,

$$\begin{aligned} \overline{B}_\lambda^L(v, t) \subseteq \overline{B}_\lambda^L(v, t) \subseteq B_{\text{Ber}}^L(v, t) \quad \text{and} \quad B_\lambda^G(v, t) \subseteq \overline{B}_\lambda^G(v, t) \subseteq B_{\text{Ber}}^G(v, t), \\ B_n^L(v, t) \subseteq \overline{B}_\lambda^L(v, t) \subseteq B_{\text{Ber}}^L(v, t) \quad \text{and} \quad B_n^G(v, t) \subseteq \overline{B}_\lambda^G(v, t) \subseteq B_{\text{Ber}}^G(v, t), \end{aligned} \quad (4.6)$$

where the last row holds when we further assume that  $W_u^{(n)} = W_u$  for all  $u \in B_n^L(v, t)$  and  $u \in B_n^G(v, t)$ , respectively, and  $B_n^L, B_n^G$  are the quantities in  $\text{BGIRG}_{W,L}(n)$ .

**4.1. Non-explosiveness with finite second moment weights.** The aim of this section is to prove Theorem 2.11 and its following corollary.

**Corollary 4.2** (Corollary to Theorem 2.11). *Let  $\mathcal{V}_{\lambda, \leq K} := \{x \in \mathcal{V}_\lambda, W_x \leq K\}$  and consider an infinite path  $\pi := (\pi_0 = v, \pi_1, \dots) \in \mathcal{E}_\lambda$ . If the total  $L$ -length of the path  $|\pi|_L = \sum_{i \geq 0} L(\pi_i, \pi_{i+1}) < \infty$ , then  $\pi \not\subseteq \mathcal{V}_{\lambda, \leq K}$ . In words, any infinite path with finite total  $L$ -length has vertices with arbitrarily large weights.*

*Proof of Corollary 4.2 subject to Theorem 2.11.* We prove the statement for  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  as well. Let us denote shortly by  $G_{\leq K}$ , (resp.,  $\overline{G}_{\leq K}$ ) the subgraph of  $\text{EGIRG}_{W,L}(\lambda)$  (resp.  $\overline{\text{EGIRG}}_{W,L}(\lambda)$ ) spanned by vertices in  $\mathcal{V}_{\lambda, \leq K}$ . Note that the statement is equivalent to showing that explosion is impossible within  $G_{\leq K}$  (resp. within  $\overline{G}_{\leq K}$ ). Since the vertex-weights  $(W_x)_{x \in \mathcal{V}_\lambda}$  are i.i.d., every  $x \in \mathcal{V}_\lambda$  belongs to the vertex set  $\mathcal{V}_{\lambda, \leq K}$  independently with probability  $\mathbb{P}(W \leq K)$ . An independent thinning of a Poisson point process (PPP) is again a PPP, thus  $\mathcal{V}_{\lambda, \leq K}$  is a PPP with intensity  $\lambda_K := \lambda \mathbb{P}(W \leq K)$ . Further, every location  $x \in \mathcal{V}_{\lambda, \leq K}$  receives an i.i.d. weight distributed as  $W_K \stackrel{d}{=} (W \mid W \leq K)$ . Finally, edges are present conditionally independently, and the edge between  $y, z \in \mathcal{V}_{\lambda, \leq K}$  with given weights  $w_y, w_z$  is present with probability  $h(y - z, w_y, w_z)$  in  $\mathcal{E}_\lambda$  (respectively,  $\overline{C}_1 \overline{g}(y - z, w_y, w_z)$  in  $\overline{\mathcal{E}}_\lambda$ ). Thus,  $G_{\leq K}$  and  $\overline{G}_{\leq K}$  can be considered as an instance of  $\text{EGIRG}_{W_K, L}(\lambda_K)$  and  $\overline{\text{EGIRG}}_{W_K, L}(\lambda_K)$ , respectively, see Definitions 2.7 and 3.1. Theorem 2.11 now applies since  $\mathbb{E}[W_K^2] < \infty$ .  $\square$

To prove the non-explosive part of Theorem 2.11, we need a general lemma. We introduce some notation related to trees first. Consider a rooted (possibly random) infinite tree  $\mathcal{T}$ . Let  $\mathcal{G}_k$ , generation  $k$ , be the vertices at graph distance  $k$  from the root.

**Lemma 4.3.** *Consider a rooted (possibly random) tree  $\mathcal{T}$ . Add i.i.d. lengths from distribution  $F_L$  to each edge with  $\mathbb{P}(L = 0) = 0$ . Let  $d_L(x, \emptyset) := \sum_{e \in P_{\emptyset, x}} L_e$  be the  $L$ -distance of  $\emptyset, x \in \mathcal{T}$ , where  $P_{\emptyset, x}$  is the unique path from vertex  $x$  to the root  $\emptyset$ . Suppose that the tree has at most exponentially growing generation sizes, that is,*

$$\exists \nu \in (1, \infty) : \mathbb{P}\left(\bigcup_{k_0 \in \mathbb{N}} \bigcap_{k \geq k_0} \{|\mathcal{G}_k| \leq \nu^k\}\right) = 1. \quad (4.7)$$

Then explosion is impossible on  $\mathcal{T}$ :

$$\mathbb{P}\left(\lim_{k \rightarrow \infty} d_L(\emptyset, \mathcal{G}_k) < \infty\right) = 0.$$

*Proof.* By (3.18), the limit  $\lim_{k \rightarrow \infty} d_L(\emptyset, \mathcal{G}_k)$  equals the explosion time of the root  $\emptyset$ . Our goal is to show that there exists a constant such that  $d_L(\emptyset, \mathcal{G}_k) \geq c_L k$  for all sufficiently large  $k$ . Hence the limit is a.s. infinite. Let us fix a small enough  $\varepsilon > 0$  and set  $t_0$  so small that  $F_L(t_0) < 1/(4\nu^2) - \varepsilon$ . Then, let us define the event

$$E_k := \{\exists y \in \mathcal{G}_k : d_L(y, \emptyset) < kt_0/2\}.$$

Then, since every vertex in  $\mathcal{G}_k$  is graph distance  $k$  from the root, by a union bound,

$$\mathbb{P}(E_k \mid |\mathcal{G}_k|) \leq |\mathcal{G}_k| F_L^{*,k}(kt_0/2), \quad (4.8)$$

where we write  $F_L^{*,k}$  for the  $k$ -fold convolution of  $L$  with itself<sup>3</sup>. Now, we bound  $F_L^{*,k}(kt_0/2)$ . Note that if more than  $k/2$  variables in the sum  $L_1 + \dots + L_k$  have length at least  $t_0$ , then the sum exceeds  $kt_0/2$ . As a result, we must have at least  $k/2$  variables with value at most  $t_0$ . Thus,

$$F_L^{*,k}(kt_0/2) \leq \binom{k}{k/2} (F_L(t_0))^{k/2} \leq 2^k ((4\nu^2)^{-1} - \varepsilon)^{k/2}. \quad (4.9)$$

Using this bound in (4.8) and that  $|\mathcal{G}_k| \leq \nu^k$  for all  $k \geq k_0$  for some  $k_0$ , we arrive to the upper bound

$$\mathbb{P}(E_k) \leq \nu^k 2^k ((4\nu^2)^{-1} - \varepsilon)^{k/2} = (1 - 4\nu^2 \varepsilon)^{k/2}.$$

<sup>3</sup>The distribution of  $L_1 + L_2 + \dots + L_k$ , with  $L_i$  i.i.d. from  $L$ .

Since the rhs is summable, the Borel-Cantelli Lemma implies that a.s. only finitely many  $E_k$  occur. That is, for all large enough  $k$ , all vertices in generation  $k$  have  $L$ -distance at least  $kt_0/2$ . Thus,  $d_L(x, \mathcal{G}_k)$  tends to infinity and the result follows.  $\square$

*Proof of Theorem 2.11, finite second moment case.* First we show non-explosion when  $\mathbb{E}[W^2] < \infty$ . We prove the statement for  $\text{EGIRG}_{W,L}(\lambda)$  and  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  as well. By translation invariance, it is enough to show that the explosion time  $Y_E^\lambda(0)$  (see Def. 2.10) of the origin is a.s. infinite, given  $0 \in \mathcal{V}_\lambda$ . Recall the convergence  $X_k = d_L^\lambda(0, \partial B_\lambda^G(0, k)) \xrightarrow{a.s.} Y_1^E(0)$  as  $k \rightarrow \infty$  from (3.18) where it is stated for  $\lambda = 1$ . Similar to the proof of Lemma 4.3, we show that for some constant  $c_L$  that depends on  $L$ ,  $X_k \geq c_L k$  for all  $k$  sufficiently large, hence the limit cannot be finite.

Consider the corresponding upper bounding  $\text{BerBRW}_\lambda(0)$ , with environment  $(x, W_x)_{x \in \mathcal{V}_\lambda}$ . By the coupling in Lemma 4.1, and (4.6),  $\partial B_\lambda^G(0, k)$  is contained in generation  $k$  (denoted by  $\mathcal{G}_k^\lambda$ ) of  $\text{BerBRW}_\lambda(0)$ , hence it is enough to show that

$$X_k^{\text{Ber}} = d_L^\lambda(0, \mathcal{G}_k^\lambda) \geq c_L k \quad (4.10)$$

for all sufficiently large  $k$ , since  $X_k \geq X_k^{\text{Ber}}$ . The inequality (4.10) follows from Lemma 4.2, when we show that  $|\mathcal{G}_k^\lambda|$  grows at most exponentially in  $k$  when  $\mathbb{E}[W^2] < \infty$ . Thus, we aim to show that when  $\mathbb{E}[W^2] < \infty$ , for some  $m < \infty$  and constant  $C$ ,

$$\mathbb{E}[|\mathcal{G}_k^\lambda|] \leq C m^k. \quad (4.11)$$

Indeed, setting  $A_k := \{|\mathcal{G}_k^\lambda| \geq (2m)^k\}$ , by Markov's inequality,  $\mathbb{P}(A_k) \leq C 2^{-k}$ , summable in  $k$ . By the Borel-Cantelli Lemma, there is a  $k_0$  such that  $A_k^c$  holds for all  $k \geq k_0$ , thus (4.7) holds with  $\nu := 2m$ . At last we show (4.11). Recall the definition of the displacement probabilities in  $\text{BerBRW}$  from (4.1). For a set  $\mathcal{A} \subset \mathbb{R}_+$ ,  $N_w^B(\mathcal{A})$ , called the *reproduction kernel*, denotes the number of children with vertex-weight in the set  $\mathcal{A}$  of an individual with vertex-weight  $w$  located at the origin. Let us fix  $s \geq 1$  an arbitrary number, and think of  $ds$  as infinitesimal, and slightly abusing notation let us write  $N_w^B([s, s+ds]) := N_w^B(ds)$ . In case of the  $\text{BerBRW}_\lambda(v)$ , the distributional identity holds:

$$N_w^B(ds) \stackrel{d}{=} \sum_{y \in \mathcal{V}_\lambda} \mathbb{1}_{\{W_y \in (s, s+ds)\}} \text{Ber}(\overline{C}_1(1 \wedge \overline{a}_1 \|y\|^{-\alpha d}(ws)^\alpha)). \quad (4.12)$$

The *expected reproduction kernel* (cf. [55, Section 5]) is defined as  $\mu(w, ds) := \mathbb{E}[N_w^B(ds)]$ , that can be bounded from above as

$$\mu(w, ds) \leq F_W(ds) \overline{C}_1 \mathbb{E} \left[ \left( \sum_{y \in \mathcal{V}_\lambda : \|y\| \leq (a_1^{1/\alpha} ws)^{1/d}} 1 \right) + \overline{a}_1 (ws)^\alpha \sum_{y \in \mathcal{V}_\lambda : \|y\| \geq a_1^{1/(\alpha d)} (ws)^{1/d}} \|y\|^{-\alpha d} \right], \quad (4.13)$$

where<sup>4</sup>  $F_L(ds) := F_L(s+ds) - F_L(s)$ . The expectation of the first sum equals  $\lambda \text{Vol}_d \cdot ws$ , with  $\text{Vol}_d$  denoting the volume of the unit ball in  $\mathbb{R}^d$ . The expectation of the second sum equals, for some constants  $c_d, \tilde{c}_d$ ,

$$(ws)^\alpha \lambda \overline{a}_1 \int_{(ws)^{1/d}}^\infty c_d t^{d-1-\alpha d} dt = (ws)^\alpha \lambda \tilde{c}_d (ws)^{1-\alpha} = \lambda \tilde{c}_d ws. \quad (4.14)$$

With  $c_\lambda := (\text{Vol}_d + \tilde{c}_d)$ , defining

$$\mu_+(w, ds) := ws F_W(ds) c_\lambda, \quad (4.15)$$

we obtain an upper bound  $\mu(w, ds) \leq \mu_+(w, ds)$  uniform in  $s, w$ .

By the translation invariance of the displacement probabilities in  $\text{BerBRW}$ , the *expected* number of individuals with given weight in generation  $k$  can be bounded from above by applying the composition operator  $*$  acting on the ‘vertex-weight’ space (type-space), that is, define

$$\mu^{*n}(w, \mathcal{A}) := \int_{\mathbb{R}_+} \mu(v, \mathcal{A}) \mu^{*(n-1)}(w, ds), \quad \mu^{*1} := \mu.$$

<sup>4</sup>The  $ds$  refers to the Lebesgue-Stieltjes integration wrt to the measure generated by  $F_L$ , e.g. when  $L$  has a density  $f_L$  then  $F_L(ds) = f_L(s)ds$ .

For instance,  $\mu^{*2}$  counts the expected number of individuals with weight in  $\mathcal{A}$  in the second generation, averaged over the environment. The rank-1 nature of the kernel  $\mu_+$  implies that its composition powers factorise. An elementary calculation using (4.15) shows that

$$\mu_+^{*k}(w, ds) = wsF_W(ds) \cdot (c_\lambda)^k \mathbb{E}[W^2]^{k-1}. \quad (4.16)$$

Since the weight of the root is distributed as  $W$ ,

$$\mathbb{E}[\mathcal{G}_k^\lambda] = \mathbb{E}[\mu^{*k}(W, \mathbb{R}_+)] \leq \mathbb{E}[\mu_+^{*k}(W, \mathbb{R}_+)] = (c_\lambda \mathbb{E}[W^2])^k \cdot \mathbb{E}[W] / \mathbb{E}[W^2]. \quad (4.17)$$

This finishes the proof with  $m := c_\lambda \mathbb{E}[W^2] < \infty$ .  $\square$

**4.2. Growth of the Bernoulli BRW when  $\tau \in (2, 3)$ .** To prove the lower bound on the  $L$ -distance in Proposition 3.6, we analyse the properties of  $\text{BerBRW}_\lambda(v)$ . We investigate the doubly-exponential growth rate of the generation sizes and the maximal displacement. Note that this analysis does not include the edge-lengths. The main result is formulated in the following proposition, that was proved for the SFP model in [51] that we extend to  $\text{EGIRG}_{W,L}(\lambda)$ :

**Proposition 4.4** (Doubly-exponential growth and maximal displacement of  $\text{BerBRW}$ ). *Consider the  $\text{BerBRW}_\lambda(v)$  in  $\mathbb{R}^d$ ,  $d \geq 1$ , with weight distribution  $W$  satisfying (2.7) with  $\tau \in (2, 3)$ , and reproduction distribution given by (4.1) with  $\alpha > 1$ . Let  $\mathcal{G}_\lambda^{\text{Ber}}(v, k)$ ,  $Z_\lambda^{\text{Ber}}(v, k)$  denote the set and size of generation  $k$  of  $\text{BerBRW}_\lambda(v)$ . We set*

$$c_k(\varepsilon, i) := 2 \exp \left\{ i(1 + \varepsilon) \left( \frac{1 + \varepsilon}{\tau - 2} \right)^k \right\}, \quad S_k(\varepsilon, i) := \exp \left\{ i \zeta \left( \frac{1 + \varepsilon}{\tau - 2} \right)^k \right\}, \quad (4.18)$$

where  $\zeta \geq \max\{2(\varepsilon/(1 + \varepsilon) + (\tau - 1)/(\tau - 2))/d, (2\alpha + \varepsilon(\tau - 2)/(1 + \varepsilon))/((\alpha - 1)d)\}$  is a constant. Recall the definition of  $B^2(v, r)$  in (2.18). Then, for every  $\varepsilon > 0$ , there exists an a.s. finite random variable  $Y(\varepsilon) < \infty$ , such that the event

$$\forall k \geq 0 : Z_\lambda^{\text{Ber}}(v, k) \leq c_k(\varepsilon, Y(\varepsilon)) \text{ and } \mathcal{G}_\lambda^{\text{Ber}}(v, k) \cap (B^2(v, S_k(\varepsilon, Y(\varepsilon))))^c = \emptyset \quad (4.19)$$

holds. Furthermore, the  $Y(\varepsilon)$  have exponentially decaying tails with

$$\mathbb{P}(Y(\varepsilon) \geq K) \leq C_\varepsilon \exp\{-\varepsilon K/4\}, \quad (4.20)$$

where the constant  $C_\varepsilon$  depends on  $\varepsilon, \tau, \alpha, \lambda, d$  and the slowly-varying function  $\ell$  only, not on  $K$ .

So, both the generation size and the maximal displacement of individuals in generation  $k$  grow at most doubly exponentially with rate  $(1 + \varepsilon)/(\tau - 2)$  and random prefactor  $Y(\varepsilon)$ . We believe based on the branching process analogue [31] that the true growth rate is  $1/(\tau - 2)$  with a random prefactor, but for our purposes, this weaker bound also suffices. It will be useful to know the definition of  $Y(\varepsilon)$  later.

**Remark 4.5.** Let us define

$$\begin{aligned} \tilde{E}_k(\varepsilon, i) &:= \{Z_\lambda^{\text{Ber}}(v, k) \leq c_k(\varepsilon, i)\} \cap \{\mathcal{G}_\lambda^{\text{Ber}}(v, k) \cap (B^2(v, S_k(\varepsilon, i)))^c = \emptyset\}, \\ H_i &:= \bigcap_{k=0}^{\infty} \tilde{E}_k(\varepsilon, i), \end{aligned} \quad (4.21)$$

that is,  $H_i$  describes the event that the generation sizes and maximal displacement of  $\text{BerBRW}_\lambda(v)$  grow (for all generations) doubly exponentially with prefactor at most  $i$  in the exponent, as in (4.18). It can be shown that  $\sum_{i=1}^{\infty} \mathbb{P}(H_i^c) < \infty$ . It follows from the Borel-Cantelli lemma that only finitely many  $H_i^c$  occur. Since  $c_k(\varepsilon, i), S_k(\varepsilon, i)$  are increasing in  $i$ ,  $H_1 \subset H_2 \dots$  is a nested sequence. Thus  $Y(\varepsilon)$  is the first  $i$  such that  $H_i$  holds. The rhs of (4.20) equals  $\sum_{i \geq K} \mathbb{P}(H_i^c)$ .

The key ingredient to proving Proposition 4.4 is Claim 4.6 below. Let us write  $N_w(ds, \geq S)$  for the number of offspring with weight in  $(s, s + ds)$  and displacement at least  $S$  of an individual with weight  $w$ , and then let  $N_w(\geq t, \geq S)$  for the number of offspring with weight at least  $t$  and displacement at least  $S$  of an individual with weight  $w$ .

**Claim 4.6** (Bounds on the number of offspring of an individual of weight  $w$ ). *Consider  $\text{BerBRW}_\lambda(v)$  from Section 4.2, with  $\tau > 2$ ,  $d \geq 1$  and  $\alpha > 1$ . Then,*

$$\begin{aligned} \mathbb{E}[N_w(\geq t, \geq 0)] &\leq M_{d,\tau,\alpha} \lambda w t^{-(\tau-2)} \ell(w), \\ \mathbb{E}[N_w(\geq 1, \geq S)] &\leq M_{d,\tau,\alpha} \lambda w^{\tau-1} S^{-(\tau-2)d} \tilde{\ell}(S^d/w) + \mathbb{1}_{\{\alpha < \tau-1\}} M_1 \lambda w^\alpha S^{-(\alpha-1)d}, \end{aligned} \quad (4.22)$$

where constants  $M_{d,\tau,\alpha}$  and  $M_1$  depend on  $d, \tau, \alpha$  and the slowly-varying function  $\ell$  only, and  $\tilde{\ell}(x) := \mathbb{1}_{\{\alpha > \tau-1\}} \ell(x) + \mathbb{1}_{\{\alpha = \tau-1\}} \int_1^x \ell(t)/t \, dt$  is slowly varying at infinity.

This claim, combined with translation invariance of the model, allows one to estimate the expected number of individuals with weights in generation-dependent intervals, which in turn also allows for estimates on the expected number of individuals with displacement at least  $S_k(\varepsilon, i)$  in generation  $k$ . These together with Markov's inequality give a summable bound on  $\mathbb{P}(H_i^c)$  (from (4.21)). As explained in Remark 4.5, a Borel-Cantelli type argument directly shows Proposition 4.4. This method is carried out in [51, Lemma 6.3] for the scale-free percolation model, based on [51, Claim 6.3], that is the same as Claim 4.6 above, modulo that  $\gamma - 1$  there is replaced by  $\tau - 2$  here. The proof of Proposition 4.4 subject to Claim 4.6 follows in the exact same way, by replacing  $\gamma - 1$  in [51] by  $\tau - 2$  here. The proof of Claim 4.6 is similar to the proof in [51, Claim 6.3], with some differences. First, the vertex set here forms a PPP while there it is  $\mathbb{Z}^d$ , and second, the exponents need to be adjusted, due to the relation between the parameters used in the SFP and GIRG models, as explained in Section 2.1.

*Proof of Claim 4.6.* We start bounding  $\mathbb{E}[N_w(\geq 1, \geq S)]$ . Due to translation invariance, we assume that the individual is located at the origin. Recall the environment is  $(x, W_x)_{x \in \mathcal{V}_\lambda}$ , and the displacement distribution from (4.1). Using the notation  $F_W(ds) = \mathbb{P}(W \in (s, s+ds))$ , for the Stieltjes integral similarly to (4.13),

$$\begin{aligned} \mathbb{E}[N_w(\geq 1, \geq S)] &\leq \overline{C}_1 \int_{1 \leq s \leq S^d/w} \mathbb{E} \left[ \sum_{x \in \mathcal{V}_\lambda: \|x\| \geq S} \overline{a}_1 \|x\|^{-\alpha d} (ws)^\alpha \right] F_W(ds) \\ &\quad + \overline{C}_1 \int_{s \geq S^d/w} \mathbb{E} \left[ \sum_{x \in \mathcal{V}_\lambda: \|x\| \leq (\overline{a}_1^{1/\alpha} ws)^{1/d}} 1 \right] F_W(ds) \\ &\quad + \overline{C}_1 \int_{s \geq S^d/w} \mathbb{E} \left[ \sum_{x \in \mathcal{V}_\lambda: \|x\| \geq (\overline{a}_1^{1/\alpha} ws)^{1/d}} \overline{a}_1 \|x\|^{-\alpha d} (ws)^\alpha \right] F_W(ds) \\ &=: T_1 + T_2 + T_3. \end{aligned} \quad (4.23)$$

where we mean the Lebesgue-Stieltjes integral wrt  $F_L$ . Bounding the expectation inside  $T_1$  and  $T_3$ , for some constant  $c_d$ , similar to (4.14), for any  $R > 0$

$$\mathbb{E} \left[ \sum_{x \in \mathcal{V}_\lambda: \|x\| \geq R} \|x\|^{-\alpha d} (ws)^\alpha \right] = (ws)^\alpha \lambda \int_R^\infty c_d t^{d-1-\alpha d} dt = (ws)^\alpha \lambda \tilde{c}_d R^{-d(\alpha-1)}. \quad (4.24)$$

Using this bound for  $R = S$  and  $R = (\overline{a}_1^{1/\alpha} ws)^{1/d}$ , for some constant  $\tilde{c}_d$ ,

$$\begin{aligned} T_1 &= \overline{C}_1 \lambda \tilde{c}_d w^\alpha S^{-d(\alpha-1)} \mathbb{E}[W^\alpha \mathbb{1}_{\{W \leq S^d/w\}}], \\ T_3 &= \overline{C}_1 \lambda \tilde{c}_d w \mathbb{E}[W \mathbb{1}_{\{W \geq S^d/w\}}]. \end{aligned} \quad (4.25)$$

The following Karamata-type theorem [15, Propositions 1.5.8, 1.5.10] yields that for any  $\beta > 0$  and  $a > 0$ , and function  $\ell$  that varies slowly at infinity,

$$\lim_{u \rightarrow \infty} \frac{1}{u^{-\beta} \ell(u)} \int_u^\infty \ell(x) x^{-\beta-1} dx = \lim_{u \rightarrow \infty} \frac{1}{u^\beta \ell(u)} \int_a^u \ell(x) x^{\beta-1} dx = \frac{1}{\beta}. \quad (4.26)$$



while for  $\beta = 0$ , the integral equals  $\ell^*(x) := \int_1^x \ell(t)/t dt$ , which is slowly varying at infinity itself and  $\ell^*(t)/\ell(t) \rightarrow \infty$  as  $t \rightarrow \infty$ .  $W$  satisfies (2.7), so we can bound the expectations in (4.25) using (4.26):

$$\mathbb{E}[W^\alpha \mathbb{1}_{\{W \leq S^d/w\}}] \leq \int_1^{S^d/w} \alpha t^{\alpha-1} \ell(t) t^{-(\tau-1)} dt \leq C_{\alpha,\tau} (S^d/w)^{\alpha-\tau+1} \ell(S^d/w), \quad (4.27)$$

when  $\alpha > \tau - 1$ , and  $\mathbb{E}[W^\alpha \mathbb{1}_{\{W \leq S^d/w\}}] \leq \mathbb{E}[W^\alpha] =: M_\alpha < \infty$  when  $\alpha < \tau - 1$ . When  $\alpha = \tau - 1$ , the integral in (4.27) itself is slowly varying, thus, with  $\ell^*(x) := \int_1^x \ell(t)/t dt$ , the bound in (4.27) remains true with  $\ell$  replaced by  $\ell^*$ . Further,

$$\mathbb{E}[W \mathbb{1}_{\{W \geq S^d/w\}}] \leq \int_{S^d/w}^\infty \ell(t) t^{-(\tau-1)} dt \leq C_\tau (S^d/w)^{-(\tau-2)} \ell(S^d/w). \quad (4.28)$$

Returning to (4.25), with  $C_{\alpha,\tau}, C_\tau$  from (4.27) and (4.28), with  $\tilde{\ell}(\cdot) := \mathbb{1}_{\{\alpha < \tau-1\}} \ell(\cdot) + \mathbb{1}_{\{\alpha = \tau-1\}} \ell^*(\cdot)$ ,

$$T_1 + T_3 \leq \bar{C}_1 \lambda \tilde{C}_d (C_{\alpha,\tau} + C_\tau) \cdot w^{\tau-1} S^{-d(\tau-2)} \tilde{\ell}(S^d/w) + \mathbb{1}_{\{\alpha < \tau-1\}} \bar{C}_1 \lambda \tilde{C}_d M_\alpha w^\alpha S^{-(\alpha-1)d}. \quad (4.29)$$

Finally, we bound  $T_2$  in (4.23). With  $\text{Vol}_d$  the volume of the unit ball in  $\mathbb{R}^d$ , as in (4.13), the expectation within  $T_2$  equals  $\text{Vol}_d \lambda \bar{a}_1^{1/\alpha} w s$ , thus, by (4.28) once more,

$$T_2 \leq \bar{C}_1 \lambda \bar{a}_1^{1/\alpha} w \mathbb{E}[W \mathbb{1}_{\{W \geq S^d/w\}}] \leq \bar{C}_1 \lambda \text{Vol}_d C_\tau \cdot w^{\tau-1} S^{-d(\tau-2)} \ell(S^d/w). \quad (4.30)$$

Combining (4.29) and (4.30) yields the desired bound in (4.22) with  $M_{d,\alpha,\tau} := \bar{C}_1 (\tilde{C}_d C_{\alpha,\tau} + \tilde{C}_d C_\tau + \text{Vol}_d C_\tau)$  and  $M_1 := \bar{C}_1 \tilde{C}_d M_\alpha$ . The bound on  $\mathbb{E}[N_w(\geq t, \geq 0)]$  follows from the calculations already done in (4.12)-(4.14), that yield, when integrating over  $s \geq t$ ,

$$\mathbb{E}[N_w(\geq t, \geq 0)] \leq \bar{C}_1 \lambda (\text{Vol}_d + \tilde{C}_d) w \int_t^\infty F_W(ds) s ds \leq M_{d,\tau,\alpha} \lambda w t^{-(\tau-2)} \ell(t). \quad (4.31)$$

where we have applied the bound in (4.28) by replacing  $S^d/w$  by  $t$ .  $\square$

## 5. LOWER BOUND ON THE WEIGHTED DISTANCE: PROOF OF PROPOSITION 3.6

In this section we prove Proposition 3.6. We start with a preliminary statement about the minimal distance between vertices in a given sized box:

**Claim 5.1.** *Consider a box  $B(H_n) := [x \pm H_n^{1/d}]^d$  in  $\tilde{\mathcal{X}}_d(n)$ . Then, in  $\text{EGIRG}_{W,L}(\lambda)$ , and in  $\text{BGIRG}_{W,L}(n)$ , for some  $c_2 > 0$ ,*

$$\mathbb{P}\left(\left\{\min_{v_1, v_2 \in B(H_n)} \|v_1 - v_2\| \leq 1/H_n^{2/d}\right\} \cup \left\{\max_{v \in B(H_n)} W_v \geq H_n^{2/(\tau-1)}\right\}\right) \leq c_2 H_n^{-3/4}. \quad (5.1)$$

*Proof.* To bound the minimal distance between two vertices in the box, one can use standard extreme value theory or simply a first moment method: given the number of vertices, the vertices are distributed uniformly within  $H_n$ , and hence the probability that there is a vertex within norm  $z > 0$  from a fixed vertex  $v \in \mathcal{V}_\lambda$ , conditioned on  $|\mathcal{V}_\lambda \cap B(H_n)|$ , is at most  $|\mathcal{V}_\lambda \cap B(H_n)| z^d \text{Vol}_d / (2^d H_n)$ , and so by Markov's inequality,

$$\begin{aligned} \mathbb{P}(\exists v_1, v_2 : \|v_1 - v_2\| \leq z) &\leq \mathbb{E}[\mathbb{P}(\exists v_1, v_2 : \|v_1 - v_2\| \leq z \mid |\mathcal{V}_\lambda \cap B(H_n)|)] \\ &\leq \mathbb{E}[|\mathcal{V}_\lambda \cap B(H_n)|^2 \cdot \text{Vol}_d z^d / (2^d H_n)] \\ &\leq (\lambda^2 H_n^2 + \lambda H_n) \text{Vol}_d z^d / (2^d H_n), \end{aligned} \quad (5.2)$$

that is at most  $\Theta(H_n^{-3/4})$  for  $z = 1/H_n^{2/d}$ . For  $\text{BGIRG}_{W,L}(n)$ , the same argument works with  $\lambda = 1$ . Bounding the maximal weight, we use that  $|B(H_n) \cap \mathcal{V}_\lambda|$  has distribution  $\text{Poi}(\lambda 2^d H_n)$ , and that each vertex has an i.i.d. weight with power-law distribution from (2.7). So, by the law of total probability, for any  $x \leq 0$ :

$$\mathbb{P}\left(\max_{v \in B(H_n) \cap \mathcal{V}_\lambda} W_v > x\right) \leq \mathbb{E}\left[\sum_{v \in B(H_n) \cap \mathcal{V}_\lambda} \mathbb{1}_{\{W_v > x\}}\right] = \ell(x) x^{-(\tau-1)} 2^d H_n \lambda. \quad (5.3)$$

Inserting  $x = H_n^{2/(\tau-1)}$  yields

$$\mathbb{P}\left(\max_{v \in B(M_n) \cap \mathcal{V}_\lambda} W_v > H_n^{2/(\tau-1)}\right) \leq \lambda \ell\left(H_n^{2/(\tau-1)}\right) H_n^{-2} 2^d H_n \leq c H_n^{-3/4} \quad (5.4)$$

where we use that  $\ell(H_n^{2/(\tau-1)}) H_n^{-1/4} < 1$  by (6.18) for large  $H_n$ . Combining the two bounds we arrive at (5.1).  $\square$

*Proof of Proposition 3.6.* Let  $q \in \{1, 2\}$  throughout the proof. We slightly abuse notation and write  $v_n^q$  for the vertices as well as their location in  $\mathbb{R}^d$ . Recall the event  $A_k$  from (2.29). First we show that the inequality (3.11) holds on  $A_{k_n}$ : Observe that on the event  $\{B_{1+\xi_n}^G(v_n^1, k_n) \cap B_{1+\xi_n}^G(v_n^2, k_n) = \emptyset\} \subseteq A_{k_n}$ , any path connecting  $v_n^1, v_n^2$  in  $\text{EGIRG}_{W,L}(1 + \xi_n)$  must intersect the boundaries of these sets, thus the bound

$$d_L^{1+\xi_n}(v_n^1, v_n^2) \geq d_L^{1+\xi_n}(v_n^1, \partial B_{1+\xi_n}^G(v_n^1, k_n)) + d_L^{1+\xi_n}(v_n^1, \partial B_{1+\xi_n}^G(v_n^2, k_n)) \quad (5.5)$$

holds<sup>5</sup>. Let us denote the shortest path connecting  $v_n^q$  to  $\partial B_{1+\xi_n}^G(v_n^q, k_n)$  in  $\text{EGIRG}_{W,L}(1 + \xi_n)$  by  $\pi_\star^q(k_n)$ . In general it can happen that the rhs of (5.5) is not a lower bound<sup>6</sup> for  $d_L(v_n^1, v_n^2)$  in  $\text{BGIRG}_{W,L}(n)$ . However, on the event  $A_{k_n}$ , this cannot happen, since on  $A_{k_n}$ , the three graphs  $B_{1+\xi_n}^G(v_n^q, k_n)$ ,  $B_n^G(v_n^1, k_n)$ ,  $B_1^G(v_n^q, k_n)$  coincide. Thus,  $\pi_\star^q(k_n)$  is *present* and also *shortest* in  $\text{BGIRG}_{W,L}(n)$  and  $\text{EGIRG}_{W,L}(1)$  as well. As a result, we arrive to (3.11). Next we bound  $\mathbb{P}(A_{k_n}^c)$  from above. We recall that, by the coupling in Claim 3.3,  $v_n^q \in \mathcal{V}_{1+\xi_n}$  for all  $n$  sufficiently large. Recall  $\epsilon_{\text{TV}}(n)$  from Assumption 2.4 and  $\epsilon(n), I_\Delta(n), I_w(n)$  from Assumption 2.5. The last two expressions are intervals for the distance of vertices as well as for vertex-weights, such that the convergence in Assumption 2.5 holds when vertex distances and vertex-weights are in these intervals. Let us thus introduce the notation for the endpoints of these intervals as follows:

$$I_\Delta(n) =: [\Delta_{\min}(n), \Delta_{\max}(n)], \quad I_w(n) =: [1, w_{\max}(n)] \quad (5.6)$$

and set:

$$H_n := \min \{ \Delta_{\min}(n)^{-d/2}, \Delta_{\max}(n)/2\sqrt{d}, w_{\max}(n)^{(\tau-1)/2}, \epsilon(n)^{-1/4}, \epsilon_{\text{TV}}(n)^{-1/4}, n^{1/(4d)} \}. \quad (5.7)$$

Note that, by the requirements in Assumption 2.5,  $I_\Delta(n) \rightarrow \infty, I_w(n) \rightarrow [1, \infty]$  implies that  $\Delta_{\min}(n) \rightarrow 0, \Delta_{\max}(n) \rightarrow \infty, w_{\max}(n) \rightarrow \infty$ , and also note that  $\epsilon(n), \epsilon_{\text{TV}}(n) \rightarrow 0$ . Hence,  $H_n \rightarrow \infty$  and it will take the role of  $H_n$  in Claim 5.1 above. Let us now define the two boxes in  $\tilde{\mathcal{X}}_d(n)$ :

$$\text{Box}(v_n^q) := [v_n^q - H_n^{1/d}, v_n^q + H_n^{1/d}]^d. \quad (5.8)$$

We divide the proof of bounding the probability of  $A_{k_n}^c$  into the following steps: we first define two events,  $E_n^{(1)}, E_n^{(2)}$ , which denote that  $\text{Box}(v_n^1)$  and  $\text{Box}(v_n^2)$  are disjoint and within  $\tilde{\mathcal{X}}_d(n)$ , and that the vertices and edges within these boxes have vertex-weights and spatial distances between vertices within  $I_w(n)$  and  $I_\Delta(n)$ , respectively. We show that these events hold whp. Then, we estimate the largest graph distance  $k_n$  from  $v_n^q$ , such that  $B_{1+\xi_n}^G(v_n^q, k_n)$  is still within  $\text{Box}(v_n^q)$ . By Lemma 4.1 and (4.6), generation  $k$  in  $\text{BerBRW}_{1+\xi_n}(v_n^q)$  is a superset of vertices at graph distance  $k$  from  $v_n^q$ . Hence, we determine the largest generation number  $k_n$  such that generation  $k_n$  in  $\text{BerBRW}_{1+\xi_n}(v_n^q), q \in \{1, 2\}$  is still within  $\text{Box}(v_n^q)$ . Proposition 4.4, bounds the maximal displacement of the generations of  $\text{BerBRW}_{1+\xi_n}(v_n^q)$  for all  $k \geq 0$ , which yields the definition of  $k_n$ .

There is a technical problem. We intend to prove our results for *two* branching random walks, within the boxes  $\text{Box}(v_n^q), q \in \{1, 2\}$ , whereas the growth estimate of one  $\text{BerBRW}$  in Proposition 4.4 depends on the whole of  $\mathbb{R}^d$ , yielding a random prefactor  $Y(\varepsilon)$ . Hence, we redefine the random prefactor  $Y(\varepsilon)$ , as used in Proposition 4.4. We cannot use this limiting  $Y(\varepsilon)$  for *two* branching random walks, one

<sup>5</sup>The distance is per definition infinite when there is no connecting path. In that case the bound is trivial.

<sup>6</sup>For instance, in  $\text{BGIRG}_{W,L}(n)$  there might be an ‘extra-edge’ from some vertex in  $B_{1+\xi_n}^G(v_n^1, k_n)$  leading to  $B_{1+\xi_n}^G(v_n^1, k_n)^c$  and the shortest path connecting  $v_n^1, v_n^2$  in  $\text{BGIRG}_{W,L}(n)$  uses this particular edge, avoiding  $\pi_\star^q(k_n)$ . Or, there is an extra edge in  $\text{BGIRG}_{W,L}(n)$  not present in  $\mathcal{E}_{1+\xi_n}$  that ‘shortcuts’  $\pi_\star^q(k_n)$  and makes an even shorter path present in  $\text{BGIRG}_{W,L}(n)$ . In these cases,  $d_L(v_n^1, v_n^2)$  might be shorter than the rhs in (5.5).

started at  $v_n^1$  and one at  $v_n^2$ , since they would not be independent. To maintain independence, we slightly modify the growth variable  $Y(\varepsilon)$  in Proposition 4.4 to be determined by the growth of the process within the box only. This way the two copies stay independent. Using these adjusted growth variables for the two BerBRWs, we find the sequence  $k_n$ , such that the first  $k_n$  generations of both BerBRWs (started at  $v_n^1, v_n^2$ ) are contained within  $\text{Box}(v_n^1), \text{Box}(v_n^2)$ , respectively.

Finally, having  $k_n$ , we show that whp the balls  $B_n^G(v_n^q, k_n), B_{1+\xi_n}^G(v_n^q, k_n), B_1^G(v_n^q, k_n)$  have identical vertex sets, edge sets and vertex-weights. Combining all of the above, we show that  $A_{k_n}$  holds whp. We now work out these steps in detail.

Let us define two events: one that these boxes are disjoint and are not intersecting the boundary of  $\tilde{\mathcal{X}}_d(n)$ , and two, that all vertex-vertex distances and vertex-weights within  $\text{BGIRG}_{W,L}(n)$  and  $\text{EGIRG}_{W,L}(1+\xi_n)$  are within  $I_\Delta(n), I_w(n)$ , respectively. Recall that  $\mathcal{V}_B(n) \subseteq \mathcal{V}_{1+\xi_n}$  from Claim 3.3. Then,

$$\begin{aligned} E_n^{(1)} &:= \{\text{Box}(v_n^1) \cap \text{Box}(v_n^2) = \emptyset, \text{Box}(v_n^q) \subseteq \tilde{\mathcal{X}}_d(n), q \in \{1, 2\}\}, \\ E_n^{(2)} &:= \{\forall v \in \mathcal{V}_{1+\xi_n} \cap (\text{Box}(v_n^1) \cup \text{Box}(v_n^2)) : W_v^{(n)} \in I_w(n), W_v \in I_w(n), \\ &\quad \forall u, v \in \mathcal{V}_{1+\xi_n} \cap (\text{Box}(v_n^1) \cup \text{Box}(v_n^2)) : \|u - v\| \in I_\Delta(n)\}. \end{aligned} \quad (5.9)$$

Recall that the position of the vertices in  $\text{BGIRG}_{W,L}(n)$  is uniform in  $\tilde{\mathcal{X}}_d(n)$  and that  $\mathcal{C}_{\max}$  has linear size. As a result, when  $v_n^1, v_n^2$  are uniformly chosen in  $\mathcal{C}_{\max}$ , their location is uniform over a linear subset of  $n$  i.i.d. position-vectors within  $\tilde{\mathcal{X}}_d(n)$ . Hence,  $\mathbb{P}(E_n^{(1)}) = 1 - e_1(n)$  for some function  $e_1(n) \downarrow 0$  as  $n \rightarrow \infty$ . Further,  $E_n^{(2)}$  captures the event that no two vertices are too close or too far away from each other, and that their weight is not too large for the bound  $\epsilon(n)$  in (2.11) in Assumption 2.5 to hold. By Claim 5.1, compared to the definition of  $H_n$  in (5.7), we see that with probability at least  $1 - c_2 H_n^{-3/4}$ :

- 1) the minimal distance between vertices within  $\text{Box}(v_n^q)$  is at most  $H_n^{-2/d} \geq \Delta_{\min}(n)$ ,
- 2) the maximal distance is at most  $\sqrt{d}2H_n \leq \Delta_{\max}$  (on the diagonal),
- 3) the maximal weight of vertices is at most  $H_n^{2/(\tau-2)} \leq w_{\max}(n)$ .

These are precisely the requirements for  $E_n^{(2)}$  to hold, hence  $\mathbb{P}(E_n^{(2)}) \geq 1 - e_2(n)$  for  $e_2(n) := c_2 H_n^{-3/4} \downarrow 0$  by Claim 5.1.

Next we turn to taking care of the independence issue and determining  $k_n$ . On  $E_n^{(1)} \cap E_n^{(2)}$ , we investigate  $\text{BerBRW}_{1+\xi_n}(v_n^q, k)$ ,  $q \in \{1, 2\}$ . The two BRWs are independent as long as they are contained in the disjoint boxes  $\text{Box}(v_n^q)$ , since the location of points in two disjoint sets in a PPP are independent, and the vertex-weights of different vertices are i.i.d. copies of  $W$ , and edges are present conditionally independently. Recall that  $\mathcal{G}_{1+\xi_n}^{\text{Ber}}(v_n^q, k)$  stands for the set of vertices in generation  $k$  of  $\text{BerBRW}_{1+\xi_n}(v_n^q)$ . Our goal is to find  $k_n$  so that the following event holds:

$$E_{n,k_n} := \left\{ \bigcap_{k \leq k_n} \{\mathcal{G}_{1+\xi_n}^{\text{Ber}}(v_n^q, k) \subseteq \text{Box}(v_n^q)\}, \quad q \in \{1, 2\} \right\}. \quad (5.10)$$

On  $E_{n,k_n}$ , Lemma 4.1 can be extended to hold for both BRWs until generation  $k_n$ , with independent environments. A combination of Claim 3.4 and Lemma 4.1, as in (4.6), yields that on  $E_{n,k_n}$ ,

$$B_{1+\xi_n}^G(v_n^q, k_n) \subseteq \text{Box}(v_n^q), \text{ and } B_{1+\xi_n}^G(v_n^q, k_n) \subseteq \tilde{\mathcal{X}}_d(n), \quad q \in \{1, 2\}, \quad (5.11)$$

see Figure 2, thus these sets with index 1, 2 are disjoint on  $E_n^{(1)}$ . Next we determine  $k_n$  such that  $E_{n,k_n}$  holds whp. We would like to apply Proposition 4.4 to say that the maximal displacement in the two BRWs grow at most doubly exponentially, as in (4.19). The random variable prefactor  $Y(\varepsilon)$ , (the double exponential growth rate of the generations) in Proposition 4.4 bounds the growth of a BRW for all generations  $k \geq 0$ . The problem with this variable is that it depends on the entire environment  $\mathcal{V}_\lambda$ . Now that we have two BRWs that only stay independent as long as they are contained in their respective boxes  $\text{Box}(v_n^q)$ , we would like to have a similar definition of doubly-exponential growth based on only observing the generations within  $\text{Box}(v_n^q)$  so that we can maintain independence. For this, we

modify the definition of  $Y(\varepsilon)$ , as in Remark 4.5, (the double exponential growth rate of the generations), to reflect this. Recall  $c_k(\varepsilon, i), S_k(\varepsilon, i)$  from (4.18). We define the event

$$\tilde{E}_k^{(q)}(\varepsilon, i) := \left\{ Z_{1+\xi_n}^{\text{Ber}}(v_n^q, k) \leq c_k(\varepsilon, i) \right\} \cap \left\{ \mathcal{G}_{1+\xi_n}^{\text{Ber}}(v_n^q, k) \cap (B_{S_k(\varepsilon, i)}^2(v_n^q))^c = \emptyset \right\}. \quad (5.12)$$

Let  $K(\varepsilon, i)$  be the largest integer  $k$  such that  $S_k(\varepsilon, i) \leq H_n^{1/d}$ , then, by (4.18),

$$K(\varepsilon, i) = \left\lfloor \frac{\log \log H_n - \log i - \log(d\zeta)}{\log((1+\varepsilon)/(\tau-2))} \right\rfloor. \quad (5.13)$$

Thus, the event that the generations sizes and displacement of  $\text{BerBRW}_\lambda(v_n^q)$  are not growing faster than doubly exponentially with prefactor  $i$  for the first  $K(\varepsilon, i)$  generations, that is, as long as the process stays within  $\text{Box}(v_n^q)$ , see (5.8), can be expressed as follows:

$$\tilde{H}_i^{(q)} := \bigcap_{k=0}^{K(\varepsilon, i)} \tilde{E}_k^{(q)}(\varepsilon, i). \quad (5.14)$$

Clearly,  $H_i \subseteq \tilde{H}_i^{(q)}$  from (4.21), so, by Remark 4.5,

$$\sum_{i=1}^{\infty} \mathbb{P}((\tilde{H}_i^{(q)})^c) \leq \sum_{i=1}^{\infty} \mathbb{P}(H_i^c) < \infty. \quad (5.15)$$

By the Borel-Cantelli lemma, we define  $\tilde{Y}_q(\varepsilon)$  as the first index  $i$  from which point on no  $(\tilde{H}_i^{(q)})^c$  occur any more. Then,  $\tilde{Y}_1(\varepsilon), \tilde{Y}_2(\varepsilon)$  are independent, since they are determined on disjoint sets of vertices. By Remark 4.5, the rhs of the tail estimate (4.20) equals  $\sum_{i \geq K} \mathbb{P}(H_i^c)$ , so, again by  $H_i \subseteq \tilde{H}_i^{(q)}$ , (4.20) is valid for  $\mathbb{P}(\tilde{Y}_q(\varepsilon) \geq K)$  as well. Let  $\omega(n) \rightarrow \infty$  be defined as

$$\omega(n) := (\log H_n)^{1/2}. \quad (5.16)$$

Then, the event

$$E_n^{(3)} := \{ \tilde{Y}_q(\varepsilon) \leq \omega(n)(\tau-2)/((1+\varepsilon)(d\zeta)), q \in \{1, 2\} \} \quad (5.17)$$

holds with probability at least  $1 - e_3(n)$  with  $e_3(n) \downarrow 0$ . An elementary calculation using (5.13) and the upper bound on  $\tilde{Y}_q(\varepsilon)$  in  $E_n^{(3)}$  yields that on  $E_n^{(3)}$ ,

$$K(\varepsilon, \tilde{Y}_q(\varepsilon)) \geq \frac{\log \log H_n}{2 \log((1+\varepsilon)/(\tau-2))} =: k_n. \quad (5.18)$$

Combining this with the definition of  $\tilde{Y}_q(\varepsilon)$  and  $\tilde{H}_i^{(q)}$ , (5.12) holds for the first  $k_n$  generations. Recall (5.8) and (5.10). Thus, on  $E_n^{(1)} \cap E_n^{(3)}$ , both  $E_{n, k_n}$  and (5.11) hold with  $k_n$  as in (5.18). To move towards the event  $A_{k_n}$  as in (2.29), we need to compare the vertex and edge-sets of  $B_n^G(v_n^q, k_n), B_{1+\xi_n}^G(v_n^q, k_n), B_1^G(v_n^q, k_n)$  and show that they are whp identical. Let

$$E_n^{(4)} := \{ |\mathcal{V}_{1+\xi_n} \cap \tilde{\mathcal{X}}_d(n)| \leq (1+2\xi_n)n, |\mathcal{V}_{1+\xi_n} \cap \text{Box}(v_n^q)| \leq (1+2\xi_n)2^d H_n, q \in \{1, 2\} \}, \quad (5.19)$$

be the event that there are not too many vertices in the boxes  $\text{Box}(v_n^q)$  and in the whole  $\tilde{\mathcal{X}}_d(n)$  that the Poisson density  $1 + \xi_n$  allows. Then, by standard calculation on Poisson variables,  $\mathbb{P}(E_n^{(4)}) = 1 - e_4(n)$  with  $e_4(n) \downarrow 0$ . By the *proof* of Claim 3.3, each vertex in  $\mathcal{V}_{1+\xi_n}$  is not in  $\mathcal{V}_B(n)$  with probability  $1 - n/|\mathcal{V}_{1+\xi_n}|$ , which is at most  $2\xi_n$  on  $E_n^{(4)}$ , and each vertex in  $\mathcal{V}_{1+\xi_n}$  is not in  $\mathcal{V}_1$  with probability  $\xi_n/(1+\xi_n) < \xi_n$ . Further, on the second event in  $E_n^{(4)}$ , the number of vertices and possible edges in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  within  $\text{Box}(v_n^q)$  is at most  $(1+2\xi_n)2^d H_n$  and  $(1+2\xi_n)^2 4^d H_n^2$ , respectively. Since on  $E_n^{(1)} \cap E_n^{(3)}$ , (5.11) holds, on  $E_n^{(1)} \cap E_n^{(3)} \cap E_n^{(4)}$ , it is an upper bound to say that in  $\overline{B}_{1+\xi_n}^G(v_n^q, k_n)$ , the number of vertices and edges until graph distance  $k_n$  is at most

$$|\mathcal{V}(\overline{B}_{1+\xi_n}^G(v_n^q, k_n))| \leq (1+2\xi_n)2^d H_n, \quad |\mathcal{E}(\overline{B}_{1+\xi_n}^G(v_n^q, k_n))| \leq (1+2\xi_n)^2 4^d H_n^2. \quad (5.20)$$

Let  $\mathcal{W}(A)$  denote the weights of vertices within a set  $A$ , and let

$$\begin{aligned} E_{n,k}^{(5)} &:= \bigcup_{q \in \{1,2\}} \{\mathcal{V}(B_n^G(v_n^q, k)) = \mathcal{V}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{V}(B_1^G(v_n^q, k)), \\ &\quad \mathcal{E}(B_n^G(v_n^q, k)) = \mathcal{E}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{E}(B_1^G(v_n^q, k)), \\ &\quad \mathcal{W}(B_n^G(v_n^q, k)) = \mathcal{W}(B_{1+\xi_n}^G(v_n^q, k)) = \mathcal{W}(B_1^G(v_n^q, k))\}. \end{aligned} \quad (5.21)$$

By the coupling in Claim 3.4, an edge in  $\mathcal{E}_{1+\xi_n}$  is in  $\mathcal{E}_1$  whenever both end vertices are in  $\mathcal{V}_1$ . Further, on  $E_n^{(2)}$  all weights and vertex-vertex distances are within  $I_\Delta(n), I_w(n)$ , and hence, conditioned that  $e$  is present in  $\bar{\mathcal{E}}_{1+\xi_n}$ , the probability that the presence of an edge in  $\mathcal{E}_{1+\xi_n}, \mathcal{E}_1$  differs from that in  $\mathcal{E}_B(n)$  is at most  $\epsilon(n) + 2\epsilon_{TV}(n)$ . This coupling also contains that the vertex-weights are identical. Finally, by (4.6), the vertex sets within graph distance  $k$  are contained in the set of individuals within generation  $k$  of the corresponding BRW. Using the bound in (5.20) on the number of vertices and potential edges, and (5.7), and  $\xi_n = \sqrt{\log n/n}$ ,

$$\begin{aligned} \mathbb{P}((E_{n,k_n}^{(5)})^c \mid E_n^{(4)} \cap E_n^{(3)} \cap E_n^{(2)} \cap E_n^{(1)}) &\leq 6\xi_n \cdot (1 + 2\xi_n)2^d H_n \\ &\quad + (\epsilon(n) + 2\epsilon_{TV}(n)) \cdot (1 + 2\xi_n)^2 4^d H_n^2 =: e_5(n) \downarrow 0, \end{aligned} \quad (5.22)$$

where the convergence of  $e_5(n)$  to 0 with  $n$  is implied by the definition of  $H_n$  in (5.7). Finally, with  $k_n$  as in (5.18), and  $A_{k_n}$  as in (2.29),  $E_n^{(1)} \cap E_n^{(2)} \cap E_n^{(3)} \cap E_n^{(4)} \cap E_{n,k_n}^{(5)} \subseteq A_{k_n}$  and hence

$$\mathbb{P}(A_{k_n}) \geq 1 - (e_1(n) + e_2(n) + e_3(n) + e_4(n) + e_5(n)) \rightarrow 1. \quad (5.23)$$

This finishes the proof.  $\square$

## 6. PERCOLATION AND BOXING METHOD ON GIRGS

In this section, we develop the necessary tools to prove Propositions 3.5 and 3.8. First, we define *weight-dependent percolation* on GIRGs and relate percolated GIRG graphs to GIRGs with modified parameters, a useful tool in bounding the length of the connecting path in Proposition 3.8. Then, we define a *boxing method* that enables to build this connecting path as well the path to infinity in Proposition 3.5. We start by defining percolation on GIRGs.

**Definition 6.1** (Weight-dependent percolation). *Consider  $\text{BGIRG}_{W,L}(n)$  or  $\text{EGIRG}_{W,L}(\lambda)$ , and a function  $p : [1, \infty)^2 \rightarrow (0, 1]$ . Conditioned on the weights of vertices, keep every existing edge  $e = (u, v)$  connecting vertices with weights  $w_u, w_v$  independently with probability  $p(w_u, w_v)$ . Let us denote the percolated graph (the graph formed by the kept edges) by  $G^p$ .*

Recall that on  $\text{BGIRG}_{W,L}(n)$ , each edge carries an i.i.d. edge-lengths  $L_e$ . Given any threshold function  $\text{thr} : [1, \infty)^2 \rightarrow (0, 1]$ , by setting

$$\mathbb{1}\{\text{edge } e = (u, v) \text{ is kept}\} = \mathbb{1}\{L_e \leq \text{thr}(w_u, w_v)\}, \quad (6.1)$$

yields a weight dependent percolation. Indeed, since the  $L_e$  are i.i.d., conditioned on the weights of vertices, edges are kept independently with probability

$$F_L(\text{thr}(w_u, w_v)) =: p(w_u, w_v),$$

which defines the function  $p$  in Definition 6.1. The next claim connects the percolated graph  $G^p$  to an instance of a GIRG with modified parameters.

**Claim 6.2.** *Consider the  $\text{BGIRG}_{W,L}(n)$ , as in Definition 2.1. Assume that  $\text{BGIRG}_{W,L}(n)$  satisfies Assumptions 2.2, Assumption 2.4 and 2.5 with  $\alpha > 1, d \geq 1, \tau > 1$ , and, for some  $c \in (0, \alpha)$  and  $\tilde{\gamma} \in (0, 1)$ , the percolation function  $p$  satisfies for all  $w_u, w_v$  that*

$$p(w_u, w_v) \geq \exp\{-c(\log w_u)^{\tilde{\gamma}} - c(\log w_v)^{\tilde{\gamma}}\}. \quad (6.2)$$

*Then the percolated graph  $G^p$  (without the edge-lengths) contains a subgraph  $\underline{G}^p$  that is an instance of a BGIRG, and the parameters of  $\underline{G}^p$  as a BGIRG again satisfy Assumptions 2.2, 2.4 and 2.5 with*

unchanged parameters  $(\alpha^P, d^P, \tau^P) = (\alpha, d, \tau)$ . The results remain valid when we percolate  $\text{EGIRG}_{W,L}(\lambda)$  instead.

*Proof.* Recall  $g_n, g_n^B$  from Definition 2.8 and from Assumption 2.5. Since edge-lengths are independent, conditioned on  $(x_i, W_i^{(n)})_{i \in [n]}$ , each edge  $e = (u, v)$  is present in  $G^P$  independently with probability

$$\begin{aligned} \mathbb{P}((u, v) \in \mathcal{E}(G^P) \mid (x_i, W_i^{(n)})_{i \in [n]}) &= g_n^B(x_u, x_v, (W_i^{(n)})_{i \in [n]}) p(W_u^{(n)}, W_v^{(n)}) \\ &\geq c_1 \left( e^{-a_2(\log W_u^{(n)})^\gamma - a_2(\log W_v^{(n)})^\gamma} \wedge \underline{a}_1 \|x_u - x_v\|^{-\alpha d} \left( W_u^{(n)} W_v^{(n)} \right)^\alpha \right) e^{-c(\log W_u^{(n)})^{\tilde{\gamma}} - c(\log W_v^{(n)})^{\tilde{\gamma}}} \\ &=: j(x_u, x_v, W_u^{(n)}, W_v^{(n)}) \end{aligned} \quad (6.3)$$

where we used that the connection probability of  $\text{BGIRG}_{W,L}(n)$  satisfies the lower bound coming from (2.16) and (2.3), as well as the lower bound on  $p$  in (6.2). Observe that formally the third row equals

$$\begin{aligned} j(x_u, x_v, W_u^{(n)}, W_v^{(n)}) &= c_1 \left( e^{-a_2(\log W_u^{(n)})^\gamma - c(\log W_u^{(n)})^{\tilde{\gamma}} - a_2(\log W_v^{(n)})^\gamma - c(\log W_v^{(n)})^{\tilde{\gamma}}} \right. \\ &\quad \left. \wedge \underline{a}_1 \|x_u - x_v\|^{-\alpha d} \left( W_u^{(n)} e^{-(c/\alpha)(\log W_u^{(n)})^{\tilde{\gamma}}} W_v^{(n)} e^{-(c/\alpha)(\log W_v^{(n)})^{\tilde{\gamma}}} \right)^\alpha \right). \end{aligned} \quad (6.4)$$

As a result, with  $m(w) := w \exp\{-(c/\alpha)(\log w)^{\tilde{\gamma}}\}$ , it is natural to define the vertex-weights for  $\underline{G}^P$  as

$$W_{p,i}^{(n)} := m(W_i^{(n)}) = W_i^{(n)} \exp\{-(c/\alpha)(\log W_i^{(n)})^{\tilde{\gamma}}\}, \quad (6.5)$$

and observe that the collection  $(W_{p,i}^{(n)})_{i \in [n]}$  is also i.i.d. In what follows we bound the function  $j(x_u, x_v, W_u^{(n)}, W_v^{(n)})$  from below by a function  $\underline{g}^P(x_u, x_v, m(W_u^{(n)}), m(W_v^{(n)}))$ , and then we will define  $\underline{G}^P$  as a BGIRG with weights from (6.5), and connection probabilities given by  $\underline{g}^P$ . To this end, we need to express the first line in (6.4) as a function of  $m(W_u^{(n)}), m(W_v^{(n)})$  instead of  $W_u^{(n)}, W_v^{(n)}$ . Since  $c \in (0, \alpha)$ , for all  $w \geq 1$ ,

$$\log(m(w)) = \log(w)(1 - (c/\alpha)(\log w)^{\tilde{\gamma}-1}) \geq (1 - c/\alpha) \log w, \quad (6.6)$$

and so we can bound,  $\gamma_{\max} := \max\{\gamma, \tilde{\gamma}\}$ ,

$$\begin{aligned} \exp\{-a_2(\log w)^\gamma - c(\log w)^{\tilde{\gamma}}\} &\geq \exp\{-(a_2 + c)(\log w)^{\gamma_{\max}}\} \\ &\geq \exp\{-(a_2 + c)(1 - c/\alpha)^{-\tilde{\gamma}}(\log m(w))^{\gamma_{\max}}\}. \end{aligned}$$

Hence, with  $\hat{c} := (a_2 + c)(1 - c/\alpha)^{-\tilde{\gamma}}$ ,

$$\begin{aligned} j(x_u, x_v, W_u^{(n)}, W_v^{(n)}) &\geq e^{-\hat{c}(\log m(W_u^{(n)}))^{\gamma_{\max}} - \hat{c}(\log m(W_v^{(n)}))^{\gamma_{\max}}} \wedge \underline{a}_1 \|x_u - x_v\|^{-\alpha d} (m(W_u^{(n)}) m(W_v^{(n)}))^\alpha \\ &=: \underline{g}^P(x_u, x_v, m(W_u^{(n)}), m(W_v^{(n)})). \end{aligned} \quad (6.7)$$

Let us hence define  $\underline{G}^P$  as the graph where the vertex-weights are i.i.d. given by (6.5) and the connection probabilities are given by  $\underline{g}^P(x_u, x_v, m(W_u^{(n)}), m(W_v^{(n)}))$ , in a coupled way to  $G^P$ , i.e. additionally thinning<sup>7</sup> some edges to obtain the right retention probabilities. Due to the conditional independence of the presence of edges, this model satisfies Definition 2.1, and the new connection probability function  $\underline{g}^P$  satisfies the bounds in Assumption 2.2 with the original  $\alpha$ . It is elementary to check that  $\underline{g}_n^P$  satisfies Assumption 2.5 since it is independent of  $n$  and hence its limit is itself. It is left to show that the distribution of  $m(W^{(n)})$  satisfies Assumption 2.4 with  $\tau^P = \tau$ . By Assumption 2.4,  $W^{(n)}$  converges in distribution to a limiting variable  $W$ . Since the map  $m(\cdot)$  is continuous,  $m(W^{(n)})$  converges in distribution to  $m(W)$ . First we show that (2.7) is still satisfied for this *limiting distribution*  $W_p := m(W)$

<sup>7</sup>This can be achieved by allocating an i.i.d. uniform  $[0, 1]$  variable to each edge and making the decision of retention based on this variable.



of  $W_p^{(n)}$ . Let  $m^{(-1)}$  denote the inverse function<sup>8</sup> of  $m(w)$ . Note that  $m(w)$  varies regularly<sup>9</sup> at infinity with index 1. As a result of [15, Theorem 1.5.12],  $m^{(-1)}$  varies regularly at infinity again with index 1 and hence can be written as  $m^{(-1)}(w) := \ell_2(w)w$  for some function  $\ell_2(w)$  that varies slowly at infinity. Using that  $W_p = m(W)$  and (2.7),

$$\mathbb{P}(W_p > x) = \mathbb{P}(W > m^{(-1)}(x)) = \ell(m^{(-1)}(x)) \cdot (m^{(-1)}(x))^{-(\tau-1)} = \ell(\ell_2(x)x)\ell_2(x)^{-(\tau-1)}x^{-(\tau-1)}. \quad (6.8)$$

Let  $\ell^*(x) := \ell(\ell_2(x)x)\ell_2(x)^{-(\tau-1)}$ . By [15, Proposition 1.5.7], it follows that  $\ell^*$  is slowly varying at infinity itself. So, the limiting distribution  $W_p$  of the weights  $(W_{p,i}^{(n)})_{i \in [n]}$  is power-law distributed with parameter  $\tau^p = \tau$ . Likewise, for the weights  $W_p^{(n)}$  themselves, using (2.6),

$$\mathbb{P}(W_p^{(n)} > x) = \mathbb{P}(W^{(n)} > m^{(-1)}(x)) = \ell^{(n)}(\ell_2(x)x)\ell_2(x)^{-(\tau-1)}x^{-(\tau-1)} =: \ell_p^{(n)}(x)x^{-(\tau-1)}. \quad (6.9)$$

Again using (2.6), we can formulate a lower and an upper bound for  $\ell_p^{(n)}$ :

$$\underline{\ell}^*(x) := \underline{\ell}(\ell_2(x)x)\ell_2(x)^{-(\tau-1)} \leq \ell_p^{(n)}(x) \leq \bar{\ell}(\ell_2(x)x)\ell_2(x)^{-(\tau-1)} =: \bar{\ell}^*(x). \quad (6.10)$$

By the same arguments as above,  $\underline{\ell}^*, \bar{\ell}^*$  vary slowly at infinity. Since  $W^{(n)} \geq W_p^{(n)}$  a.s.,

$$\mathbb{P}(W_p^{(n)} > M_n) \leq \mathbb{P}(W^{(n)} > M_n) = o(n^{-1}). \quad (6.11)$$

Finally,  $m(1) = 1$  hence  $W_p^{(n)} \geq 1$  also holds. Hence, the weights  $(W_{p,i}^{(n)})_{i \in [n]}$  satisfy Assumption 2.4. The modification to  $\text{EGIRG}_{W,L}(\lambda)$  is analogous by working with  $h$  from (2.14) instead of  $\tilde{g}_n$ .  $\square$

In the rest of the section we introduce a boxing structure that will be useful in establishing the connecting path in Proposition 3.8 as well as the existence of  $\mathcal{C}_\lambda^\infty$  in Proposition 3.5. For constants  $C, D > 1$ , to be defined shortly, and a parameter  $\mu$ , let us define the following:

$$D_k := \mu^{DC^k/d}, \quad R_k := \mu^{C^k/d}, \quad (6.12)$$

where  $\mu$  can tend to infinity. We define a *boxing system* centred at  $u \in \mathbb{R}^d$ , by defining for  $k \geq 0$ ,

$$\begin{aligned} \text{Box}_k(u) &:= \{x \in \mathbb{R}^d : \|x - u\|_\infty \leq D_k/2\}, \\ \Gamma_k(u) &:= \text{Box}_k(u) \setminus \text{Box}_{k-1}(u), \quad k \geq 1, \quad \Gamma_0(u) := \text{Box}_0(u), \\ B_k(z) &:= \{x \in \mathbb{R}^d : \|x - z\|_\infty \leq R_k/2\}. \end{aligned} \quad (6.13)$$

Let us further set  $k_{\max} = k_{\max}(\mu, n) := \max\{k \in \mathbb{N} : D_k^d \leq n\}$ . We ‘tile’ the annuli  $\Gamma_k(u)$  with disjoint boxes of radius  $R_k$ , that we call subboxes, in a natural way. We order the subboxes arbitrarily and refer to the  $i^{\text{th}}$  subbox in  $\Gamma_k(u)$  as  $B_k^{(i)}$  and denote the number of subboxes in  $\Gamma_k(u)$  by  $b_k$ . For sufficiently large  $\mu$ , by the ratio of volumes it can be shown that

$$\mu^{(D-1)C^k}/2 \leq b_k \leq \mu^{(D-1)C^k}. \quad (6.14)$$

Within each subbox  $B_k^{(i)}$ , let  $c_k^{(i)}$  be the vertex with the highest weight, i.e.  $c_k^{(i)} := \arg \max_{v \in B_k^{(i)}} \{W_v\}$ , that we call the ‘centre’ of  $B_k^{(i)}$ . The idea is to find a path of centres, one in each annulus, with a total  $L$ -length that is small, as displayed in Figures 3 and 4. Let us fix a small  $\varepsilon > 0$  and define

$$\begin{aligned} \delta(\varepsilon) &:= (\tau - 2)\varepsilon / (2(\tau - 1)), \quad C(\varepsilon) := (1 - \varepsilon) / (\tau - 2), \\ D(\varepsilon) &:= (1 - \delta(\varepsilon)) \frac{1 - \varepsilon / (\tau - 1)}{1 - \varepsilon} - \delta(\varepsilon) / 2. \end{aligned} \quad (6.15)$$

It is elementary to show that  $D(\varepsilon) \geq 1$ . In the following lemma, we give bounds on the weight of centres, and show that centres are connected to many centres in the next annulus.

<sup>8</sup>The function  $m$  is monotone on  $[x_0, \infty)$ , for some  $x_0 > 0$ , thus its inverse can be defined on this interval, which is sufficient since the assumption 2.4 is concerned with the tail behaviour of the weight distribution.

<sup>9</sup>A function  $z(x)$  varies regularly at infinity with index  $\varrho$  when  $\lim_{x \rightarrow \infty} z(cx)/z(x) = c^\varrho$  for all  $c > 0$ .

**Lemma 6.3** (Weights and subgraph of centres). *Consider  $\text{EGIRG}_{W,L}(\lambda)$  (resp.  $\text{BGIRG}_{W,L}(n)$ ) with parameters  $d \geq 1$ ,  $\tau \in (2, 3)$ ,  $\alpha > 1$ . Let  $u \in \mathbb{R}^d$ , and consider the boxing with parameters  $\mu > 1$ , fixed small  $\varepsilon > 0$ , and let  $C := C(\varepsilon)$ ,  $D := D(\varepsilon)$ ,  $\delta := \delta(\varepsilon)$ . Define  $N_j(c_k^{(i)})$  as the number of centres in  $\Gamma_j(u)$  that are connected to  $c_k^{(i)}$  and define the events*

$$\begin{aligned} F_k^{(1)} &:= \left\{ \forall i \leq b_k : W_{c_k^{(i)}} \in \left[ \mu^{C^k(1-\delta)/(\tau-1)}, \mu^{C^k(1+\delta)/(\tau-1)} \right] \right\}, \\ F_k^{(2)} &:= \left\{ \forall i \leq b_k : N_{k+1}(c_k^{(i)}), N_k(c_k^{(i)}) \geq \exp\{C^k(D-1)(\log \mu)/2\}/2 \right\}. \end{aligned} \quad (6.16)$$

Then, for some  $c_3(\varepsilon) > 0$ , whenever  $\mu > \mu_0(\varepsilon)$ ,

$$1 - \mathbb{P}(\cap_{k \geq 0} (F_k^{(1)} \cap F_k^{(2)})) \leq c_3(\mu^{-\delta/4} + \exp\{-\lambda \mu^{\delta/2}/2\} + \exp\{-\mu^{(D-1)C/2}/32\}). \quad (6.17)$$

Finally, for a vertex  $v$  in  $\text{Box}_0(u)$ , conditioned on  $\cap_{k \geq 0} F_k^{(1)} \cap \{W_v = \mu\}$ ,  $v$  is connected to at least one centre in annulus  $\Gamma_1(v)$  with probability at least  $1 - \exp\{-\mu^{(D-1)C/2}/16\}$ .

The statements in Lemma 6.3 are quenched, in the sense that they hold for almost all realisations of  $\text{EGIRG}_{W,L}(\lambda)$  or  $\text{BGIRG}_{W,L}(n)$ , depending in which model you apply the lemma. The message of the lemma is the following: a centre of a subbox is connected to many centres of subboxes in the next annulus. By following paths that jump from centres to centres in annuli with increasing size, it is possible to create paths going to infinity. This result is an important factor in various proofs: it shows that there is an infinite component in  $\text{EGIRG}_{W,L}(\lambda)$ . Moreover, since the degrees of centres grow doubly exponentially, it is even possible to construct paths going to infinity in a way that the total edge-length of the constructed path will be small (by always choosing the shortest option), which leads to connecting the shortest paths between two vertices. The idea of the proof of Lemma 6.3 is that the boxes and subboxes are defined such that the weights of centres are sufficiently large and they are sufficiently close to the centres in consecutive annuli, and therefore the connection probability between these centres is fairly large. More specifically, it can be bounded by the first argument of  $g$ , as in (2.3).

*Proof of Proposition 3.5 subject to Lemma 6.3.* Using the boxing structure as defined in (6.13) and the connectivity argument in Lemma 6.3, we prove the existence of an infinite component. Fix a sequence  $(\mu_1 < \mu_2 < \dots)$  so that the error bound in (6.17) is summable for the sequence  $\mu_i$ . Construct a boxing system centred at the origin with parameter  $\mu_i$ , for each  $i \geq 1$ . Since the error probabilities are summable, the Borel-Cantelli Lemma implies that there is an index  $i_0$  for which  $\cap_{k \geq 0} (F_k^{(1)} \cap F_k^{(2)})$  will hold with parameter  $\mu_{i_0}$ . Thus, in this boxing system, an infinite component exists since we can simply find an infinite path along a sequence of centres  $c_1^{(i_1)}, c_2^{(i_2)}, \dots$ , in annuli with increasing index. This component is almost surely unique by [65, Theorem 6.3], which proves Proposition 3.5.  $\square$

*Proof of Lemma 6.3.* For this proof we make use of Potter's theorem [15, Theorem 1.5.6] that states that for any  $\varepsilon > 0$ ,

$$\lim_{x \rightarrow \infty} \ell(x) x^{-\varepsilon} = 0, \quad \lim_{x \rightarrow \infty} \ell(x) x^\varepsilon = \infty. \quad (6.18)$$

We prove the statements in Lemma 6.3 for the  $\text{EGIRG}_{W,L}(\lambda)$  model and discuss how to adjust the proof for the  $\text{BGIRG}_{W,L}(n)$  model. We start by estimating the weights of the centres. Using that the vertices are generated by a PPP with intensity  $\lambda$ ,  $\text{Vol}(B_k^{(i)}) = R_k^d$ , the power-law probability in (2.7), and conditioning on  $|B_k^{(i)} \cap \mathcal{V}_\lambda|$  yields

$$\begin{aligned} \mathbb{P}\left(\max_{v \in B_k^{(i)} \cap \mathcal{V}_\lambda} W_v \leq x\right) &= \sum_{j=0}^{\infty} \mathbb{P}(W \leq x)^j e^{-R_k^d \lambda} (R_k^d \lambda)^j / j! \\ &= \exp\left\{-R_k^d \lambda + R_k^d \lambda (1 - \ell(x) x^{-(\tau-1)})\right\} \\ &= \exp\left\{-R_k^d \lambda \ell(x) x^{-(\tau-1)}\right\}. \end{aligned} \quad (6.19)$$

Inserting the lower bound in  $F_k^{(1)}$  in (6.16) for  $x$  and using the definition of  $R_k$  from (6.12), we find that  $R_k^d x^{-(\tau-1)} = \mu^{\delta C^k}$  and hence

$$\mathbb{P}\left(\max_{v \in B_k^{(i)} \cap \mathcal{V}_\lambda} W_v \leq \mu^{C^k(1-\delta)/(\tau-1)}\right) \leq \exp\left\{-\lambda \ell(\mu^{C^k(1-\delta)/(\tau-1)}) \mu^{\delta C^k}\right\} \leq \exp\left\{-\lambda \mu^{(\delta/2)C^k}\right\}, \quad (6.20)$$

where we use that  $\ell(\mu^{(1-\delta)/(\tau-1)C^k}) \mu^{(\delta/2)C^k} > 1$  for large  $\mu$  by (6.18). For the upper bound, we use the first moment method to find

$$\mathbb{P}\left(\max_{v \in B_k^{(i)} \cap \mathcal{V}_\lambda} W_v > x\right) \leq \mathbb{E}\left[\sum_{v \in B_k^{(i)} \cap \mathcal{V}_\lambda} \mathbb{1}_{\{W_v > x\}}\right] = \ell(x) x^{-(\tau-1)} R_k^d \lambda. \quad (6.21)$$

Inserting the upper bound from (6.16) for  $x$  yields

$$\mathbb{P}\left(\max_{v \in B_k^{(i)}} W_v > \mu^{C^k(1+\delta)/(\tau-1)}\right) \leq \ell\left(\mu^{C^k(1+\delta)/(\tau-1)}\right) \lambda \mu^{-\delta C^k} \leq \lambda \mu^{-(3/4)\delta C^k}, \quad (6.22)$$

where we use that  $\ell(\mu^{C^k(1+\delta)/(\tau-1)}) \mu^{-(\delta/4)C^k} < 1$  by (6.18) for large  $\mu$ . Combining the bounds in (6.20) and (6.22) with a union bound over  $k$  and the at most  $b_k$  subboxes, for some constant  $c_3 > 0$ ,

$$\begin{aligned} \mathbb{P}\left(\left(\cap_{k \geq 0} F_k^{(1)}\right)^c\right) &\leq \sum_{k=0}^{\infty} b_k (\lambda \mu^{-(3/4)\delta C^k} + \exp\{-\lambda \mu^{(\delta/2)C^k}\}) \\ &\leq c_3 \mu^{(D-1-(3/4)\delta)} + c_3 \exp\{-\lambda \mu^{\delta/2}/2\}, \end{aligned} \quad (6.23)$$

where we can further use that  $D-1-(3/4)\delta < -\delta/4$ , as follows from the definition of  $D(\varepsilon), \delta(\varepsilon)$  in (6.15), and the bound in (6.14). This yields the first two terms on the rhs of (6.17).

We now turn to the events  $(F_k^{(2)})_{k \in \mathbb{N}}$ . The heuristics of the proof are the following: consider a centre  $c_k^{(i)}$  of a subbox in annulus  $k$ . We would like to estimate the number of centres in annulus  $k+1$  that it is connected to,  $N_{k+1}(c_k^{(i)})$ . For this, on  $F_k^{(1)} \cap F_{k+1}^{(1)}$ , we have good bounds on the weight of  $c_k^{(i)}$  and the centres in annulus  $k+1$ . Their Euclidean distance is at most the diameter of the outer annulus. Since the connection probabilities are bounded from below by the function  $g$  from (2.3), we show that the connection probability is large, in the sense that the minimum is taken by its first term when applied to  $c_k^{(i)}$  and a centre in annulus  $k$ . Since the presence of edges is conditionally independent given the weights involved,  $N_{k+1}(c_k^{(i)})$  is bounded from below by a binomial random variable with the number of centres in annulus  $k+1$  and the lower bound we derived on the connection probability as its parameters. Concentration of this binomial variable then yields the desired lower bound. Now we work out the details. Recall Claim 2.6 and (2.3), in particular that the connection probability in EGIRG with vertices of weight  $w_1, w_2$  at distance  $\|x\|$  away is at least

$$\underline{c}_1 g(x, w_1, w_2) = \underline{c}_1 \left( e^{-a_2((\log w_1)^\gamma + (\log w_2)^\gamma)} \wedge \underline{a}_1 \|x\|^{-\alpha d} (w_1 w_2)^\alpha \right) \quad (6.24)$$

We condition that the event  $F_k^{(1)} \cap F_{k+1}^{(1)}$  holds. For two centres in annuli  $k, k+1$ , respectively, their weights satisfy the bounds in  $F_k^{(1)} \cap F_{k+1}^{(1)}$ , and their Euclidean distance is at most  $\sqrt{d}D_{k+1}$  by (6.12) and (6.13), hence, the second term in the minimum in (6.24) is at least

$$a_1 \|c_k^{(i)} - c_{k+1}^{(j)}\|^{-\alpha d} (W_{c_k^{(i)}} W_{c_{k+1}^{(j)}})^\alpha \geq a_1 d^{-\alpha d/2} \mu^{\alpha(C^k + C^{k+1})(1-\delta)/(\tau-1)} / (D_{k+1})^{\alpha d} =: a_{1,d} \mu^{\alpha Z C^k} \quad (6.25)$$

with  $Z := (1-\delta)(1+C)/(\tau-1) - CD$ . It can be shown using (6.15) that  $Z = \varepsilon(1-\varepsilon)/4(\tau-1) > 0$ . Let us abbreviate  $l(w) := \exp\{-a_2(\log w)^\gamma\}$ . Then, the first term in (6.24) is, conditioned on the event  $F_k^{(1)} \cap F_{k+1}^{(1)}$ , using the lower bound in (6.16),

$$l(W_{c_k^{(i)}}) l(W_{c_{k+1}^{(j)}}) \geq \exp\left\{-a_2 \left(\frac{1+\delta}{\tau-1}\right)^\gamma \left(\log \mu^{C^k}\right)^\gamma (1+C^\gamma)\right\} =: a_{\mu,k} / \underline{c}_1, \quad (6.26)$$

with  $c_1$  the same constant as in Claim 2.6. Hence,

$$\mathbb{P}\left(c_k^{(i)} \leftrightarrow c_{k+1}^{(j)} \mid \cap_{k \geq 1} F_k^{(1)}\right) \geq c_1 \min \left\{ a_{\mu,k}/c_1, a_{1,d} \mu^{\alpha_Z C^k} \right\} = a_{\mu,k} \quad (6.27)$$

since the second term in the minimum is larger than its first term when  $\mu > \mu_0$  is sufficiently large. Recall that the number of subboxes in annulus  $k+1$  is  $b_{k+1}$  (see before (6.14)). Since conditionally on the weights of vertices, edges are present independently,

$$\left( N_{k+1}(c_k^{(i)}) \mid W_{c_k^{(i)}}, (W_{c_{k+1}^{(j)}})_{j \leq b_{k+1}} \right) \stackrel{d}{\geq} \text{Bin}(b_{k+1}, a_{\mu,k}), \quad (6.28)$$

with  $\text{Bin}(n, p)$  a binomial rv with parameters  $n$  and  $p$ , where the lhs denotes the rv  $N_{k+1}(c_k^{(i)})$  conditionally on the weights of vertices involved, that should satisfy the event  $\cap_{k \geq 1} F_k^{(1)}$ . By the concentration of binomial rvs [47, Theorem 2.21],

$$\mathbb{P}(\text{Bin}(n, p) \leq np/2) \leq e^{-np/8}. \quad (6.29)$$

Combining this with (6.28), (6.27) and the bound on  $b_{k+1}$  in (6.14) yields

$$\mathbb{P}\left(N_{k+1}(c_k^{(i)}) \leq b_{k+1} a_{\mu,k}/2 \mid \cap_{k \geq 1} F_k^{(1)}\right) \leq \exp\{-a_{\mu,k} b_{k+1}/8\} \leq \exp\left\{-\mu^{(D-1)C^{k+1}/2}/16\right\}, \quad (6.30)$$

for  $\mu \geq \mu_1 \geq \mu_0$  sufficiently large, since  $\gamma \in (0, 1)$  in  $a_{\mu,k}$  and the bound on  $b_{k+1}$  in (6.14). Elementary calculation yields that  $a_{\mu,k} b_{k+1}/2 \geq \exp\{C^{k+1}(D-1)(\log \mu)/2\}/2$ , as in (6.16). A union bound results in

$$\mathbb{P}\left(\cap_{k \geq 1} F_k^{(2)} \mid \cap_{k \geq 1} F_k^{(1)}\right) \leq \sum_{k=0}^{\infty} b_k \exp\left\{-\mu^{(D-1)C^{k+1}/2}/16\right\} \leq c_3 \exp\left\{-\mu^{(D-1)C/2}/32\right\}, \quad (6.31)$$

where the final statement follows since the sum is dominated by a geometric series, and increasing  $c_3 > 0$  when necessary. The bound on  $N_k(c_k^{(i)})$  in (6.16) follows in a similar manner, where now set  $a_{\mu,k} := c_1 l(\mu^{C^k(1+\delta)/(\tau-1)})^2$ . The proof of the last statement in Lemma 6.3 follows by noting that  $v$  can take the role of  $c_0^{(i)}$ , i.e. the bound on the second event  $F_0^{(2)}$  for  $k=0$  directly implies the estimate on  $N_1(c_0^{(i)}) > 1$ , when  $\mu > \mu_0$  sufficiently large so that  $\exp\{(\log \mu)(D-1)C/2\} > 1$ .

For the results to hold in the  $\text{BGIRG}_{W,L}(n)$  model, the following adjustments are necessary: First note that the events  $F_k^{(1)}$  and  $F_k^{(2)}$  are only relevant for  $k \leq k_{\max}(\mu, n) := \max\{k \in \mathbb{N} : D_k^d \leq n\}$ , so that the annulus  $\Gamma_{k_{\max}}$  has a non-trivial intersection with  $\tilde{\mathcal{X}}_d(n)$ . Further, in  $\text{BGIRG}_{W,L}(n)$ , the weight distribution depends on  $n$ , as in Assumption 2.4, hence, all slowly-varying functions  $\ell$  in the proof above should be changed to either  $\underline{\ell}$  or  $\bar{\ell}$ . In particular, in the proof of the lower bound on the weights  $\underline{\ell}$  should be used, turning the second line of (6.19) into an inequality. In the upper bound for the weights,  $\bar{\ell}$  can be used, and then (6.22) remains valid. The replacement of  $\ell$  with  $\underline{\ell}$  or  $\bar{\ell}$  could be done without leaving the interval  $[0, M_n]$  where these bounds hold, since  $M_n \gg n^{1/(\tau-1)}$  by the condition in Assumption 2.4, while the maximal weight vertex in the whole graph  $\text{BGIRG}_{W,L}(n)$  - following the calculation in (6.19) with deterministically  $n$  vertices - is of order  $n^{1/\tau-1}$ . In (6.19), the Poisson probabilities are replaced by

$$\mathbb{P}(|B_k^{(i)} \cap \mathcal{V}(\text{BGIRG}_{W,L}(n))| = j) = \binom{n}{j} (\text{Vol}(B_k^{(i)})/n)^j (1 - \text{Vol}(B_k^{(i)})/n)^{n-j}, \quad (6.32)$$

but the results in the upper bounds in (6.19) and (6.22) stay valid with  $\lambda = 1$ .  $\square$

## 7. BEST EXPLOSIVE PATH CAN BE FOLLOWED

In this section we prove Lemma 3.7. We decompose its proof into three claims that we first state all at once and then prove one-by-one.

**Claim 7.1.** *Let  $E_{7.1}(n)$  be the event that both uniformly chosen vertices  $v_n^1, v_n^2 \in \mathcal{C}_{\max}$  of  $\text{BGIRG}_{W,L}(n)$  are also in  $\mathcal{C}_{\infty}^1$ , the infinite component of  $\text{EGIRG}_{W,L}(1)$ . Then  $E_{7.1}(n)$  holds with probability  $1 - e_{7.1}(n)$ , for some  $e_{7.1}(n) \downarrow 0$ .*

**Claim 7.2.** *The infinite component of  $\text{EGIRG}_{W,L}(\lambda)$  is a.s. explosive. That is, conditioned on  $0 \in \mathcal{C}_\infty^\lambda$  there exists an infinite path  $\pi = \{0 = \pi_0, \pi_1, \dots\}$  such that  $(\pi_i, \pi_{i+1}) \in \mathcal{E}(\text{EGIRG}_{W,L}(\lambda))$  and  $\sum_{i \geq 0} L_{(\pi_i, \pi_{i+1})} < \infty$ .*

We denote the infinite path from  $v$  with shortest total  $L$ -length in  $\text{EGIRG}_{W,L}(\lambda)$  by  $\pi_{\text{opt}}^\lambda(v)$ , defined in (2.21). Claim 7.2 implies that  $|\pi_{\text{opt}}^\lambda(v)|_L < \infty$  a.s. By Corollary 4.2, a.s., every infinite path with finite total  $L$ -length contains vertices with arbitrary large weights. This is important since we would like to find large-weight vertices on  $\pi_{\text{opt}}^1(v_n^1), \pi_{\text{opt}}^1(v_n^2)$  in order to construct a connecting path and hence upper bound the distance between these vertices.

**Claim 7.3.** *On the event  $E_{7.1}(n)$  from Claim 7.1, let  $\tilde{v}_n^q(K)$  denote the first vertex on  $\pi_{\text{opt}}^1(v_n^q)$  with weight at least  $K$ . Let  $E_{7.3}(n, K) \subset E_{7.1}(n)$  denote the event that the segment  $\pi_{\text{opt}}^1[v_n^q, \tilde{v}_n^q(K)]$  is within  $\tilde{\mathcal{X}}_d(n)$  and also contained in  $\text{BGIRG}_{W,L}(n)$ . Then,  $\mathbb{P}(E_{7.3}(n, K)) \geq 1 - e_{7.3}(n, K)$  for some function  $e_{7.3}(K, n)$  with  $\lim_{n \rightarrow \infty} e_{7.3}(n, K) = 0$  for all fixed  $K$ .*

Before we proceed to the proofs, we quickly prove Lemma 3.7 subject to these three claims.

*Proof of Lemma 3.7 subject to Claims 7.1 - 7.3.* By noting that  $\tilde{A}_{n,K} := E_{7.3}(n, K)$ , the proof follows by Claims 7.1 - 7.3 with  $f(n, K) := e_{7.1}(n) + e_{7.3}(n, K)$ .  $\square$

*Proof of Theorem 2.11 when  $\tau \in (2, 3)$  subject to Claim 7.2.* Suppose first that  $\mathbf{I}(L) = \infty$ . In this case, due to the coupling in Lemma 4.1 and (4.6), non-explosion of  $\text{BerBRW}_\lambda(v)$  implies non-explosion of  $\text{EGIRG}_{W,L}(\lambda)$ . Non-explosion of  $\text{BerBRW}_\lambda(v)$  is the content of the proof of [51, Theorem 1.1] that can be found in [51, Section 7]. The idea is that when  $\mathbf{I}(L) = \infty$  and the generation sizes of a BRW grow at most doubly exponentially, the sum over  $k$  of the length of the *shortest edge* between generation  $k$  and  $k+1$  is already infinite a.s., hence, any path to infinity must also have infinite total length a.s. We refer the reader to [51, Section 7] for further details. The proof of the explosive direction follows from Claim 7.2.  $\square$

*Proof of Claim 7.1.* We distinguish two cases, whether  $W_{v_n^q}^{(n)} \leq \max\{\xi_n, \epsilon_{\text{TV}}(n), \epsilon(n)\}^{-1/(2(\tau+\eta_1))} =: w(n)$  or not, with  $\eta_1 > 0$  small. We start with the first case. We make use of a result by Bringmann et al., [22, Lemma 5.5]. Fix a weight  $w \geq 1$  that might depend on  $n$  later and a small constant  $\eta_1 > 0$ . Let the event  $E_n^1(v, w)$  denote that a vertex  $v \in [n]$  with weight at most  $w$  has component size larger than  $w^{\tau+\eta_1}$ , and  $E_n^2(v, w)$  denote that vertex  $v$  is within graph distance  $w^{\tau+\eta_1}$  of some vertex with weight at least  $w$ . Then, by [22, Lemma 5.5], if  $v$  is chosen uniformly from the vertices with weight  $\leq w$ ,

$$\mathbb{P}(E_n^1(v, w) \cap (E_n^2(v, w))^c) \leq \exp\{-w^{\eta_2}\}, \quad (7.1)$$

for some  $\eta_2 = \eta_2(\eta_1) > 0$ . Note that if  $v_n^q$  is uniformly chosen in  $\mathcal{C}_{\max}$  and we assume that  $W_{v_n^q}^{(n)} \leq w$ , then, since the weights are originally i.i.d.,  $v_n^q$  is also uniformly chosen in  $\mathcal{C}_{\max} \cap \{v \in [n] : W_v^{(n)} \leq w\}$ . Since  $|\mathcal{C}_{\max}|$  is linear in  $n$  by [22, Theorem 5.3], it follows that for any  $w$ ,  $\mathbb{P}(E_n^2(v, w) \mid v \in \mathcal{C}_{\max}) \geq 1 - c_4 \exp\{-w^{\eta_2}\}$  for some  $c_4 > 0$ . On  $E_n^2(v, w)$ , let us choose a vertex  $v_\geq$  that has weight at least  $w$  and graph distance at most  $w^{(\tau+\eta_1)}$  from  $v$ , denote its weight by  $\mu_\geq$  and denote the path from  $v$  to  $v_\geq$  by  $\pi[v, v_\geq]$ . Using the coupling in Claim 3.3, each vertex in  $\mathcal{V}_B(n)$  is not in  $\mathcal{V}_{1-\xi_n} \subseteq \mathcal{V}_1$  with probability  $\mathbb{E}[1 - |\mathcal{V}_{1-\xi_n} \cap \tilde{\mathcal{X}}_d(n)|/n] = \xi_n$ . Given that a vertex  $u$  is in both models, we can couple its weight in  $\text{BGIRG}_{W,L}(n)$  and in  $\text{EGIRG}_{W,L}(1 - \xi_n)$  with error  $\epsilon_{\text{TV}}(n)$ . Finally, by (3.8) in Claim 3.4, given that vertex-weights are equal, each edge that is part of  $\mathcal{E}_B(n)$  is not part of  $\mathcal{E}_{1-\xi_n} \subseteq \mathcal{E}_1$  with probability at most  $\epsilon(n)$ . So, setting  $w = w(n)$ ,

$$\begin{aligned} \mathbb{P}(\pi[v, v_\geq] \not\subseteq \text{EGIRG}_{W,L}(1) \mid E_n^2(v, w)) &\leq \mathbb{P}(\exists u \in \pi[v, v_\geq] : u \notin \mathcal{V}_{1-\xi_n}) \\ &\quad + \mathbb{P}(\exists u \in \pi[v, v_\geq] : W_u^{(n)} \neq W_u \mid \forall u \in \pi[v, v_\geq] : u \in \mathcal{V}_{1-\xi_n}) \\ &\quad + \mathbb{P}(\exists (u_1, u_2) \in \mathcal{E}(\pi[v, v_\geq]) : (u_1, u_2) \notin \mathcal{E}_{1-\xi_n} \mid \forall u \in \pi[v, v_\geq] : u \in \mathcal{V}_{1-\xi_n}, W_u^{(n)} = W_u) \\ &\leq w(n)^{\tau+\eta_1} (\xi_n + \epsilon_{\text{TV}}(n) + \epsilon(n)), \end{aligned} \quad (7.2)$$

that tends to zero by the choice of  $w(n)$ . To summarise, assuming that a vertex  $v$  is in the giant component of  $\text{BGIRG}_{W,L}(n)$ , with high probability we find a vertex  $v_{\geq}$  with weight at least  $w(n)$  that is connected to  $v$ , and the whole segment of the path between  $v, v_{\geq}$  is part of  $\text{EGIRG}_{W,L}(1)$  whp as well. Next,  $v_{\geq}$  will serve as the initial vertex to construct a component with infinitely many vertices  $\text{EGIRG}_{W,L}(1)$ . Namely, we construct the boxing system  $\text{EGIRG}_{W,L}(1)$ , described in (6.12)–(6.13), and centred at  $v_{\geq}$  with starting weight  $\mu := \mu_{\geq} \geq w(n)$ . Lemma 6.3, then ensures that the centres of subboxes in each annulus of this boxing system have sufficiently many connections to centres in the next annulus, in particular, at least one connection. Starting from  $v_{\geq}$ , we find that it is connected to a centre in annulus one, and iteratively following an arbitrary centre in the next annulus that the centre is connected to, we find an infinite path  $\{v_{\geq} = c_0^{(i_0)}, c_1^{(i_1)}, \dots\}$  containing a centre from each annulus  $\Gamma_k, k \geq 1$ .

This construction fails with probability that is the rhs of (6.17), that is at most  $c_5 \mu_{\geq}(n)^{-\delta(\varepsilon)/4} \leq c_5 w(n)^{-\delta(\varepsilon)/4}$ , for some  $c_5 > c_3$ , since the dominating term is the first one in (6.17) when  $\mu = \mu(n) \rightarrow \infty$ . As a result,  $v$  is *not* part of an infinite component of  $\text{EGIRG}_{W,L}(1)$  with probability at most

$$e_{7.1}(n) := c_4 \exp\{-\max\{\xi_n, \epsilon_{\text{TV}}(n), \epsilon(n)\}^{-\eta_2/2(\tau+\eta_1)}\} \\ + 2 \max\{\xi_n, \epsilon_{\text{TV}}(n), \epsilon(n)\}^{1/2} + c_5 \max\{\xi_n, \epsilon_{\text{TV}}(n), \epsilon(n)\}^{\delta(\varepsilon)/8}, \quad (7.3)$$

which tends to zero as  $n \rightarrow \infty$ . Since the infinite component of  $\text{EGIRG}_{W,L}(1)$  is almost surely unique, this is sufficient. When the weight of  $v_n^q, W_{v_n^q}^{(n)} \geq w(n)$ , then we can directly construct a boxing system in  $\text{EGIRG}_{W,L}(1)$  around  $v_n^q$  and use Lemma 6.3, i.e.  $v_n^q = v_{\geq}$  can be set in the proof above. So, (7.3) is an upper bound in this case as well, finishing the proof.  $\square$

*Proof of Claim 7.2.* It is enough to construct an infinite path starting at  $0 \in \mathcal{C}_{\infty}^{\lambda}$  with a.s. finite total length. Let  $(\mu_i)_{i \in \mathbb{N}}$  be an increasing sequence such that the rhs of (6.17) with  $\mu = \mu_i$  are summable over  $i \geq 1$ . Then, construct a boxing system around 0 for each  $i \geq 1$  and let  $i_0$  be the first index such that  $\cap_{k \geq 0} (F_k^{(1)} \cap F_k^{(2)})$  holds for the boxing system with parameter  $\mu_{i_0}$  (see (6.13)). The rv  $i_0$  is a.s. finite by the Borel-Cantelli lemma. Consider this boxing system. Let  $c_0^{(1)}$  be the vertex with maximal weight in  $\Gamma_0$ . Then, since  $\cap_{k \geq 0} (F_k^{(1)} \cap F_k^{(2)})$  holds, there is an infinite path through centres in consecutive annuli starting from  $c_0^{(1)}$ , implying that  $c_0^{(1)} \in \mathcal{C}_{\infty}^{\lambda}$  holds. By the uniqueness of  $\mathcal{C}_{\infty}^{\lambda}$  and the fact that  $\|c_0^{(1)}\|_{\infty} \leq \mu_{i_0}$ , the graph distance  $d_G(0, c_0^{(1)})$  is a proper random variable, and so is the  $L$ -length of the path  $\pi[0, c_0^{(1)}]$  realising the graph distance. We iteratively continue this path into a path  $\pi_{\text{greedy}}$  in a greedy manner so that it has finite total  $L$ -length, as follows: let us call  $c_0^{(1)} =: c_0^{\text{gr}} \in \pi_{\text{greedy}}$ , and for  $k \geq 0$ , from  $c_k^{\text{gr}} \in \pi_{\text{greedy}}$ , take the minimal-length edge among those leading to the  $N_{k+1}(c_k^{\text{gr}})$  many centres in  $\Gamma_{k+1}$  that are connected to  $c_k^{\text{gr}}$  and let  $c_{k+1}^{\text{gr}}$  be its other end-vertex. By the lower bound on  $N_{k+1}(c_k^{(i)})$  on  $F_k^{(2)}$  in (6.16), the total length of  $\pi_{\text{greedy}}$  is, in distribution, bounded from above by

$$|\pi_{\text{greedy}}|_L \stackrel{d}{\leq} |\pi[0, c_0^{(1)}]|_L + \sum_{k=1}^{\infty} \min\{L_{k,1}, L_{k,2}, \dots, L_{k,d_k}\}, \quad (7.4)$$

with  $d_k := \exp\{C^k(D-1)(\log \mu_{i_0})/2\}/2$ , which is doubly-exponential in  $k$ . It follows from [51, Lemma 2.5] that the integral criterion in (2.22) is equivalent to the a.s. convergence of the sum in (7.4).  $\square$

*Proof of Claim 7.3.* In this proof, we make use of the event  $A_{k_n}$ , as in (2.29). Recall  $k_n$  from (5.18), where  $H_n$  is defined in (5.7), and that on  $A_{k_n}$ , the graph distance balls  $B_n^G(v_n^q, k_n), B_1^G(v_n^q, k_n)$  as *weighted graphs* coincide, and are part of  $\tilde{\mathcal{X}}_d(n)$  for  $q \in \{1, 2\}$ , as well as that this event holds with probability at least  $1 - \sum_{i \leq 5} e_i(n) \rightarrow 1$  by (5.23). Recall also  $E_{7.1}$ , the event that  $v_n^q \in \mathcal{C}_{\max}$  is in the infinite component of  $\text{EGIRG}_{W,L}(1)$ . On  $E_{7.1}$ , the shortest explosive path  $\pi_{\text{opt}}^1(v_n^q)$  exists. Let  $G_K(v_n^q)$  denote the number of edges in the segment  $\pi_{\text{opt}}^1[v_n^q, \hat{v}_n^q(K)]$ . We define the event

$$E_{n,K}^{(6)} := \{G_K(v_n^q) < k_n\} \quad (7.5)$$



On  $E_{n,K}^{(6)} \cap E_{7.1} \cap A_{k_n}$ ,  $\pi_{\text{opt}}^1[v_n^q, \tilde{v}_n^q(K)] \subseteq \text{BGIRG}_{W,L}(n)$  holds. Generally

$$\begin{aligned} \mathbb{P}((A \cap B \cap C)^c) &\leq \mathbb{P}(A^c | B \cap C) \mathbb{P}(B \cap C) + \mathbb{P}(A^c | (B \cap C)^c) \mathbb{P}((B \cap C)^c) + \mathbb{P}(B^c) + \mathbb{P}(C^c) \\ &\leq \mathbb{P}(A^c | B \cap C) + 2\mathbb{P}(B^c) + 2\mathbb{P}(C^c). \end{aligned}$$

With  $A := E_{n,K}^{(6)}$ ,  $B := E_{7.1}$ ,  $C := A_{k_n}$ , this estimate turns into

$$\mathbb{P}(\pi_{\text{opt}}^1[v_n^q, \tilde{v}_n^q(K)] \not\subseteq \text{BGIRG}_{W,L}(n)) \leq 2\mathbb{P}((E_{7.1})^c) + 2\mathbb{P}(A_{k_n}^c) + \mathbb{P}((E_{n,K}^{(6)})^c | E_{7.1} \cap A_{k_n}). \quad (7.6)$$

Note that on the conditioning, the graph distance balls with radius  $k_n \rightarrow \infty$  coincide in  $\text{BGIRG}_{W,L}(n)$  and  $\text{EGIRG}_{W,L}(1)$  as weighted graphs. Thus, by the translation invariance of  $\text{EGIRG}_{W,L}(1)$ , on  $A_{k_n}$ ,  $G_K(v_n^q) \leq k_n$  has the same distribution as  $G_K(0) \leq k_n$  conditioned on  $0 \in \mathcal{C}_\infty^1$ , i.e. the dependence of  $G_K(v_n^q)$  on  $n$  is not apparent. Thus the proof is finished with

$$e_{7.3}(n, K) := 2e_{7.1}(n) + \sum_{i \leq 5} 2e_i(n) + \mathbb{P}(G_K(0) \geq k_n | 0 \in \mathcal{C}_\infty^1). \quad (7.7)$$

□

*Proof of Proposition 3.8.* By Prop. 3.5 and Claim 7.2, the infinite component of  $\mathcal{C}_\infty^1$  in  $\text{EGIRG}_{W,L}(1)$  is explosive, and by Claim 7.1,  $v_n^q \in \mathcal{C}_\infty^1$ ,  $q \in \{1, 2\}$  whp. Hence,  $\pi_{\text{opt}}^1(v_n^q)$ , the optimal infinite ray from  $v_n^q$  in  $\text{EGIRG}_{W,L}(1)$  with finite total  $L$ -length exists whp (see (2.21)). Recall  $\tilde{v}_n^q(K)$  as the first vertex on  $\pi_{\text{opt}}^1(v_n^q)$  with weight exceeding  $K > 0$ , that exists by Corollary 4.2, and that  $T_K(v_n^q) = d_L^1(v_n^q, \tilde{v}_n^q(K))$ . On the event  $\tilde{A}_{n,K}$  in (3.12), the segment  $\pi_{\text{opt}}^1[v_n^q, \tilde{v}_n^q(K)]$  is present in  $\text{BGIRG}_{W,L}(n)$ , and  $\tilde{A}_{n,K}$  holds with probability at least  $1 - f(n, K)$  by Lemma 3.7. We shortly write  $\tilde{v}_n^q := \tilde{v}_n^q(K)$  from now on in this proof.

Our aim is to prove (3.13), i.e. that on  $\tilde{A}_{n,K}$  we can connect  $\tilde{v}_1^n$  to  $\tilde{v}_2^n$  with a short path within  $\text{BGIRG}_{W,L}(n)$ . We establish this connection *in the percolated graph*  $G^{\text{thr}}$ . More precisely, we do the following: Fix  $c \leq \alpha$  and any  $\tilde{\gamma} \in (0, 1)$ , and set the threshold function to be

$$\text{thr}(w_1, w_2) := F_L^{(-1)}\left(\exp\{-c(\log w_u)^{\tilde{\gamma}} - c(\log w_v)^{\tilde{\gamma}}\}\right). \quad (7.8)$$

Following the argumentation in (6.1), we keep each edge  $e = (u, v)$  if and only if its edge-weight  $L_e \leq \text{thr}(w_u, w_v)$ . This gives us a graph  $G^{\text{thr}}$ , that is the outcome of a weight-dependent percolation on  $\text{BGIRG}_{W,L}(n)$ , as in Definition 6.1, with

$$p(w_u, w_v) = F_L(\text{thr}(w_u, w_v)) \geq \exp\{-c(\log w_u)^{\tilde{\gamma}} - c(\log w_v)^{\tilde{\gamma}}\}. \quad (7.9)$$

The last inequality holds by the right-continuity of the distribution function  $F_L$  and the definition of  $F_L^{(-1)}$ . Hence, the conditions of Claim 6.2 are satisfied with this choice of  $\text{thr}$ , and  $G^{\text{thr}}$  contains a subgraph  $\underline{G}^{\text{p}}$  that is an instance of a BGIRG, that still satisfies Assumptions 2.2, 2.4 and 2.5. Recall from the proof of Claim 6.2, from (6.5), that in  $\underline{G}^{\text{p}}$  a vertex has weight  $W_{\text{p},i}^{(n)} = m(W_i^{(n)})$ . Recall also that the function  $m$  is increasing eventually, in particular, we assume that  $K$  is such that it is increasing on  $[K, \infty)$ .

Now we return to  $\tilde{v}_n^q$ . For  $q \in \{1, 2\}$ , we construct two boxing systems as in (6.12)–(6.13), centred at  $\tilde{v}_n^q$  with parameter  $\mu^{(q)} := m(W_{\tilde{v}_n^q}^{(n)}) \geq m(K)$ ,  $q \in \{1, 2\}$ , respectively, as shown in Figure 3. We apply Lemma 6.3 to the boxing system *within the subgraph of the percolated graph*  $\underline{G}^{\text{p}}$ . This is an important point of the argument, since then we have a deterministic upper bound (given by  $\text{thr}$ ) on the edge-weight of an edge that is part of  $\underline{G}^{\text{p}}$ . Then, with the error probabilities given in Lemma 6.3, each subbox centre has at least one connection within  $\underline{G}^{\text{p}}$  to a centre in the next annulus, and hence we can create a path  $\pi^q$  from  $\tilde{v}_n^q$  through the centres of sub-boxes in consecutive annuli, as in Figure 4. We would like to connect two such paths by connecting a centre in one of the systems to a centre in the other system. This may not be directly possible, hence we *merge* the two boxing systems in what follows.

Without loss of generality, we assume  $\mu^{(1)} \leq \mu^{(2)}$ . We refer to the boxing system around  $\tilde{v}_n^q$  as System- $q$  and denote all radii and structures as in (6.12) and (6.13) and the centres  $c_k^{(i)}$  with a superscript or argument  $q$  to show which system they belong to. Recall  $C = C(\varepsilon)$  from (6.15) and set

$$r := \lceil (\log \log \mu^{(2)} - \log \log \mu^{(1)}) / \log C \rceil = (\log \log \mu^{(2)} - \log \log \mu^{(1)}) / \log C + a_{1,2}, \quad (7.10)$$

with  $a_{1,2} \in [0, 1)$  an upper fractional part, hence  $r \geq 1$ . In words,  $r$  is the first index  $k$  when the annulus  $k$  in System-1 becomes larger than annulus 1 in System-2. Let us modify System-1 as follows: define  $(\text{Box}_k^1, \Gamma_k^1)_{k \leq r-1}$  as given by (6.13) with parameter  $\mu^{(1)}$ , and then let, for  $j \geq 0$ ,

$$\text{Box}_{r+j}^1 := \{x \in \mathbb{R}^d : \|x - \tilde{v}_n^1\| \leq D_j^{(2)}\}, \quad (7.11)$$

i.e. these boxes have radius  $D_j^{(2)} = (\mu^{(2)})^{DC^j/d}$  instead of  $D_{r+j}^{(1)} = (\mu^{(1)})^{DC^{r+j}} = (\mu^{(2)})^{DC^{j+a_{1,2}}/d}$ , and, since  $a_{1,2} \geq 0$ ,  $D_r^{(1)} > D_0^{(2)}$ . In words, this means that we *shrink* the size of the annuli with index at least  $r$  somewhat, to have the same base  $\mu^{(2)}$  as System-2. Note that  $\text{Box}_r^{(1)}$  is still bigger than  $\text{Box}_{r-1}^{(1)}$ . As usual,  $\Gamma_i^{(1)} := \text{Box}_i^{(1)} \setminus \text{Box}_{i-1}^{(1)}$ . Let us further use subboxes of diameter  $R_0^{(2)} = (\mu^{(2)})^{1/d}$  instead of  $R_r^{(1)}$  in annulus  $\Gamma_r^{(1)}$ . We claim that the error probabilities in (6.17) remain valid in Lemma 6.3 with  $\mu$  replaced by  $m(K)$  for the modified System-1. For this it is enough to investigate  $\mathbb{P}(F_r^{(1)}(\tilde{v}_n^1))$  and  $\mathbb{P}(F_{r-1}^{(2)}(\tilde{v}_n^1) \mid \cap_{k \geq 1} F_k^{(1)}(\tilde{v}_n^1))$ , since for the other indices, the error estimates remain valid with  $\mu$  either replaced by  $\mu^{(1)}$  or  $\mu^{(2)}$ , both at least  $m(K)$ .

Using the definition of  $r$ , we rewrite  $F_{r-1}^{(1)}(\tilde{v}_n^1)$  in terms of  $\mu^{(2)}$ , and modify  $F_r^{(1)}(\tilde{v}_n^1)$  to accommodate the fact that the subbox sizes have volume  $\mu^{(2)}$  as follows, with  $\delta = \delta(\varepsilon)$  as in (6.15):

$$\begin{aligned} F_{r-1}^{(1)}(\tilde{v}_n^1) &= \left\{ W_{c_{r-1}^{(i)}(1)}^{(n)} \in \left[ (\mu^{(2)})^{(1-\delta)/(\tau-1)C^{-(1-a_{1,2})}}, (\mu^{(2)})^{(1+\delta)/(\tau-1)C^{-(1-a_{1,2})}} \right] \right\}, \\ \tilde{F}_r^{(1)}(\tilde{v}_n^1) &:= \left\{ W_{c_r^{(i)}(1)}^{(n)} \in \left[ (\mu^{(2)})^{(1-\delta)/(\tau-1)}, (\mu^{(2)})^{(1+\delta)/(\tau-1)} \right] \right\}. \end{aligned} \quad (7.12)$$

Then, the estimate on  $F_0^{(1)}(\tilde{v}_n^2)$  with parameter  $\mu^{(2)}$  is valid for  $\mathbb{P}(\tilde{F}_r^{(1)}(\tilde{v}_n^1))$ . Further, analogously to the arguments between (6.25) - (6.27), the connection probabilities of individual centres are minimised by the first argument, i.e. with  $l(w) = \exp\{-\hat{c}(\log w)^{\gamma^{\max}}\}$ , for some  $\hat{c} > 0$  and some  $\gamma^{\max} \in (0, 1)$ ,

$$\mathbb{P}\left(c_{r-1}^{(i)}(1) \leftrightarrow c_r^{(j)}(1) \mid F_{r-1}^{(1)} \cap \tilde{F}_r^{(1)}\right) \geq c_1 l\left((\mu^{(2)})^{(1+\delta)/(\tau-1)}\right) l\left((\mu^{(2)})^{(1+\delta)/(\tau-1)C^{-(1-a_{1,2})}}\right) =: \tilde{a}_{\mu,r}, \quad (7.13)$$

Let us write  $\tilde{b}_r^1$  for the number of subboxes in  $\Gamma_r^1$ . Then, by (6.13) and (7.11),

$$\tilde{b}_r^1 = ((\mu^{(2)})^D - (\mu^{(2)})^{DC^{-(1-a_{1,2})}}) / \mu^{(2)} \geq (\mu^{(2)})^{D-1}/4 \quad (7.14)$$

for all sufficiently large  $K$ , since  $\mu^{(2)} > m(K)$ . It might happen that (some of) the annuli are not completely within  $\tilde{\mathcal{X}}_d(n)$ . This happens when either of the  $\tilde{v}_n^q$  is close to the boundary of  $\tilde{\mathcal{X}}_d(n)$ . However, Lemma 6.3 still applies since only at most half of the volume of each annulus is outside  $\tilde{\mathcal{X}}_d(n)$  and therefore the part of the annulus  $\Gamma_k^q, k \geq 0$  that lies within  $\tilde{\mathcal{X}}_d(n)$  still grows doubly exponentially with  $k$ , so the lower bound on  $b_k$  in (6.14) could be modified by a factor 1/2, and the results of Lemma 6.3 still hold. Recall  $N_{k+1}(c_k^{(i)})$  from Lemma 6.3. Then, analogous to (6.28) - (6.30), by the independence of the presence of edges conditioned on the weights,

$$\begin{aligned} \left( N_r(c_{r-1}^{(i)}(1)) \mid W_{c_r^{(i)}}^{(n)}, (W_{c_r^{(j)}}^{(n)})_{j \leq \tilde{b}_r^1} \right) &\stackrel{d}{\geq} \text{Bin}(\tilde{b}_r^1, \tilde{a}_{\mu,r}), \\ \mathbb{P}\left(N_r(c_{r-1}^{(i)}(1)) \leq \tilde{b}_r^1 \tilde{a}_{\mu,r} / 2\right) &\leq 2 \exp\left\{-\tilde{b}_r^1 \tilde{a}_{\mu,r} / 8\right\} \leq 2 \exp\left\{-(\mu^{(2)})^{(D-1)/2} / 16\right\}, \end{aligned} \quad (7.15)$$

where the second inequality follows for sufficiently large  $\mu^{(2)}$  (and thus  $K$ ), using the lower bound on  $\tilde{b}_r^1$  and bounding  $\tilde{a}_{\mu,r}$  from below by using that  $\gamma^{\max} \in (0, 1)$ . Finally, noting that  $\tilde{b}_r^1 \tilde{a}_{\mu,r} \geq (\mu^{(2)})^{(D-1)/2} / 2 = \exp\{(D-1)(\log \mu^{(2)})/2\}$  for all sufficiently large  $\mu^{(2)}$  (and thus  $K$ ) finishes the proof of the statement.

To conclude, we can build paths from both vertices  $\tilde{v}_n^1(K), \tilde{v}_n^1(K)$  through centres of boxes in the modified boxing system, with error probability as in Lemma 6.3 with  $\mu$  replaced by  $m(K)$ . Next we *connect* these paths. We introduce

$$k^* := \max\{k \in \mathbb{N} \mid D_k^{(2)} < \|\tilde{v}_n^1 - \tilde{v}_n^2\|/(2\sqrt{d})\}, \quad (7.16)$$

the largest (random)  $k$  such that  $\text{Box}_{r+k}^{(1)}$  and  $\text{Box}_k^{(2)}$  are disjoint. Note that per definition, the next boxes, with index  $k^* + 1, r + k^* + 1$ , respectively, would be centred around  $\tilde{v}_n^2, \tilde{v}_n^1$ . Instead, we now define  $\text{Box}_{k^*+1}^{(1,2)}$  as a box *centred at*  $0 \in \mathbb{R}^d$ , containing *both*  $\text{Box}_{r+k^*}^{(1)}(\tilde{v}_n^1)$  and  $\text{Box}_{k^*}^{(2)}(\tilde{v}_n^2)$ . So, we let

$$\begin{aligned} D_{k^*+1} &:= \left( (\sqrt{d}/2 + 1)(\mu^{(2)})^{DC^{k^*+1}} \right)^{1/d}, & R_{k^*+1} &:= \left( (\mu^{(2)})^{C^{k^*+1}} \right)^{1/d}, \\ \text{Box}_{k^*+1}^{(1,2)} &:= \{x \in \mathbb{R}^d \mid \|x\|_\infty \leq D_{k^*+1}/2\}, \\ \Gamma_{k^*+1}^{(1,2)} &:= \text{Box}_{k^*+1}^{(1,2)}(\tilde{v}_n^1, \tilde{v}_n^2) \setminus \left( \text{Box}_{m+k^*}^{(1)} \cup \text{Box}_{k^*}^{(2)} \right). \end{aligned} \quad (7.17)$$

Observe that the change is not in the exponent, just in the prefactor of the size of the box, hence, the error bounds on  $F_{r+k^*}^{(2)}(\tilde{v}_n^1)$  and  $F_{k^*}^{(2)}(\tilde{v}_n^2)$  in Lemma 6.3 hold in System-1 and 2, respectively. Thus, with error probability twice the rhs of (6.17), there exist two paths *in the percolated graph*  $\underline{G}^p \subseteq G^{\text{thr}}$

$$\pi^{(1)} := \{\tilde{v}_n^1, c_1^{(i_1)}(1), c_2^{(i_2)}(1), \dots, c_{r+k^*+1}^{(i_{r+k^*+1})}(1)\}, \quad \pi^{(2)} := \{\tilde{v}_n^2, c_1^{(j_1)}(2), c_2^{(j_2)}(2), \dots, c_{k^*+1}^{(j_{k^*+1})}(2)\}. \quad (7.18)$$

through the annuli  $\Gamma_k^{(q)}(\tilde{v}_n^q)$ . Note that there are  $b_{k^*+1} \geq (\mu^{(2)})^{(D-1)C^{k^*+1}}/2$  many centres in  $\Gamma_{k^*+1}^{(1,2)}$ . The connection probability between any two centres within  $\Gamma_{k^*+1}^{(1,2)}$  is at least

$$\mathbb{P}\left(c_{k^*+1}^{(i)} \leftrightarrow c_{k^*+1}^{(j)} \mid F_{k^*+1}^{(1)}\right) \geq c_1 l \left( (\mu^{(2)})^{(1+\delta)/(\tau-1)C^{k^*+1}} \right)^2 \geq c_1 (\mu^{(2)})^{-2a_2 \left( (1+\delta)/(\tau-1)C^{k^*+1} \right)^\gamma} =: a_{\mu^{(2)}}, \quad (7.19)$$

where  $F_{k^*+1}^{(1)}$  is the event defined in (6.16), as in the proof of Lemma 6.3. When  $c_{r+k^*+1}^{(i_{r+k^*+1})}(1) = c_{k^*+1}^{(j_{k^*+1})}(2)$  or there is an edge between these two vertices in  $G^p$ , a connection is established. If these are not the case, then we find a third centre that connects them. Since we condition on the environment, edges are present independently with probability at least  $a_{\mu^{(2)}}$ . Hence,

$$\begin{aligned} \mathbb{P}(\nexists i \leq b_{k^*+1}, c_{r+k^*+1}^{(i_{r+k^*+1})}(1) \leftrightarrow c_{k^*+1}^{(i)} \leftrightarrow c_{k^*+1}^{(j_{k^*+1})}(2) \mid F_{k^*+1}^{(1)}) &\leq (1 - a_{\mu^{(2)}})^{2b_{k^*+1}} \\ &\leq \exp\{-2b_{k^*+1}a_{\mu^{(2)}}\} \leq \exp\{-(\mu^{(2)})^{(D-1)C^{k^*+1}}/2\} \end{aligned} \quad (7.20)$$

for all large enough  $\mu^{(2)}$ . As a result, this error term can be merged in with the other error terms in Lemma 6.3. Let us denote this connecting piece (of either 0, 1 or 2 edges) by  $\pi^{\text{conn}}$ . What is left is to bound the  $L$ -length of the constructed path. For this we use the definition of  $\text{thr}$  in (7.8) and its monotonicity in  $w_u, w_v$ , and the fact that we know a lower bound on the weights of the centres along the constructed path, due to the events  $\cap_{k \geq 1} (F_k^{(1)}(\tilde{v}_n^1) \cap F_k^{(1)}(\tilde{v}_n^2))$ . Observe that the first  $r$  vertices of the path  $\pi^{(1)}$  are in System-1 while the rest is in the merged system, with base  $\mu^{(2)}$ . All combined yields

$$\begin{aligned} |\pi^{(1)}|_L + |\pi^{(2)}|_L + |\pi^{\text{conn}}|_L &\leq \\ &\sum_{k=0}^r F_L^{(-1)} \left( \exp\{-c \log((\mu^{(1)})^{(1-\delta)/(\tau-1)C^k})^{\tilde{\gamma}} - c \log((\mu^{(1)})^{(1-\delta)/(\tau-1)C^{k+1}})^{\tilde{\gamma}}\} \right) \\ &\quad + 2 \sum_{k=0}^{k^*+1} F_L^{(-1)} \left( \exp\{-c \log((\mu^{(2)})^{(1-\delta)/(\tau-1)C^k})^{\tilde{\gamma}} - c \log((\mu^{(2)})^{(1-\delta)/(\tau-1)C^{k+1}})^{\tilde{\gamma}}\} \right) \\ &\leq 3 \sum_{k=0}^{\infty} F_L^{(-1)} \left( \exp\left\{-C^{\gamma k} ((1-\delta)/(\tau-1) \log K)^{\tilde{\gamma}}\right\} \right) =: \varepsilon_K, \end{aligned} \quad (7.21)$$

where we have used that  $\mu^{(q)} \geq m(K) \geq K^{1/2}$ ,  $q \in \{1, 2\}$ , for all sufficiently large  $K$  in the last inequality. The integrability condition in (2.22) is equivalent to the summability of the final expression in (7.21) by [51, Claim 4.5]. Further,  $\varepsilon_K \rightarrow 0$  as  $K \rightarrow \infty$ . Thus, set  $\varepsilon_K$  as in (7.21) and error probability  $\eta(K)$  as twice the rhs of (6.17), with  $\mu$  replaced by  $m(K) \geq K^{1/2}$ . This finishes the proof.  $\square$

## 8. EXTENSION TO SCALE FREE PERCOLATION

In this section, we give a sketch of the proof of Theorem 2.14. The proof of Theorem 2.14 is somewhat simpler than that of Theorem 2.12, since the model is already defined on an infinite space, using translation invariant connection probabilities. As a result, there is no need to extend the underlying space, and there is also no need to couple the model to one with limiting connection probabilities.

*Proof of Theorem 2.14.* In [51, Theorem 1.7], a lower bound for the  $L$ -distance is already proved, stating that

$$\lim_{m \rightarrow \infty} d_L(0, \lfloor m\varepsilon \rfloor) - Y^S(0) - Y^S(\lfloor m\varepsilon \rfloor) \geq 0. \quad (8.1)$$

The proof is analogous to the proof of Proposition 3.6, i.e. it uses a BRW upper bound to bound the maximal displacement of the clusters around 0 and  $\lfloor m\varepsilon \rfloor$ , and a lower bound as (3.15), the sum of the shortest path leading to the boundary of vertices within graph distance  $k_n$ . Then, it can again be shown that these variables tend to the explosion time of the two vertices, as was done between (3.15) – (3.19).

To establish the corresponding upper bound, we should connect the two shortest explosive paths starting from 0,  $\lfloor m\varepsilon \rfloor$  in  $\text{SFP}_{W,L}$  by an arbitrarily short path. This is the missing direction from [51, Theorem 1.7]. To do this, we use weight-dependent percolation as for GIRG. The connection probabilities of vertices in  $\text{SFP}_{W,L}$  are too specific to be able to say that the percolated graph – as in Definition 6.1 – is still a scale-free percolation model. Hence, let us generalise the connection probability in (2.5) in a similar manner as for GIRG. Instead of (2.5), let us assume that, for some function  $h_{\text{SFP}}$ ,

$$\mathbb{P}(u \leftrightarrow v \mid \|u - v\| \geq 1, (W_i)_{i \in \mathbb{Z}^d}) = h_{\text{SFP}}(\|u - v\|, W_u, W_v), \quad (8.2)$$

that satisfies for some  $\tilde{\alpha} > d, \lambda > 0$ , and with  $l(w) = \exp\{-a_2 \log(w)^\gamma\}$  with  $a_2 > 0, \gamma \in (0, 1)$ ,

$$\underline{c}_1 \left( l(w_u)l(w_v) \wedge \lambda \|u - v\|^{-\tilde{\alpha}} w_u w_v \right) \leq h_{\text{SFP}}(\|u - v\|, w_u, w_v) \leq \overline{C}_1 \left( 1 \wedge \lambda \|u - v\|^{-\tilde{\alpha}} w_u w_v \right), \quad (8.3)$$

for some constants  $\underline{c}_1, \overline{C}_1 \in (0, 1]$ . Note that with this new assumption, Claim 6.2 remains valid: whenever we apply weight-dependent percolation to non-nearest neighbour edges of an SFP model satisfying (8.3), with percolation function  $p(w_u, w_v)$  satisfying the inequality in (6.2), the obtained percolated subgraph  $G^p$  contains  $\underline{G}^p$ , an instance of a scale-free percolation model again satisfying (8.3), with the same parameters  $\tilde{\alpha}, \tilde{\tau}$  as before. Further, the boxing method developed in Lemma 6.3 remains valid for the SFP model under (8.3), since we only use the upper and lower bound on the conditional probabilities of edges on the environment. The proof even becomes easier since the number of vertices in each subbox is deterministic in SFP.

We argue that a result similar to Proposition 3.8 is valid for SFP as well. For SFP, the event  $\tilde{A}_{n,K}$  holds with probability 1, since it is not necessary to couple the model to its limit. Let us set  $0 := v^1, \lfloor m\varepsilon \rfloor := v^2$ . Similar to the proof for GIRG, we denote by  $\pi_{\text{opt}}^q$  the shortest infinite path emanating from  $v^q$ . We fix a  $K > 0$  and denote by  $\tilde{v}^q(K)$  the first vertex on  $\pi_{\text{opt}}^q$  with weight at least  $K$ ,  $q \in \{1, 2\}$ . These vertices exist a.s. since the subgraph spanned by vertices with weight  $< K$  cannot be explosive, see Corollary 4.2. Then, as in the proof of Proposition 3.8, we apply weight-dependent percolation on  $\text{SFP}_{W,L}$  with threshold function  $\text{thr}(w_u, w_v)$  given in (7.8), and use the edge-lengths  $L_e$  to realise this percolation, as in (6.1). This yields  $p(w_u, w_v)$  as in (7.9) above, satisfying the conditions of Claim 6.2. So, the percolated graph  $G^{\text{thr}}$  contains a subgraph  $\underline{G}^p$  that is still an SFP, with new edge weights  $W^p = W \exp\{-c \log(W)^\gamma\}$ , i.e. here  $\alpha = 1$ .

We connect the two vertices  $\tilde{v}^1(K), \tilde{v}^2(K)$  by a path of length at most  $\varepsilon_K$  in  $\underline{G}^p \subseteq G^{\text{thr}}$ . For this we can word-by-word follow the proof of Proposition 3.8 in Section 7: we build two boxing systems, centred around  $\tilde{v}^q(K)$  with parameter  $\mu^{(q)} := W_{\tilde{v}^q(K)}^p \geq m(K)$ . When  $\mu^{(1)} \leq \mu^{(2)}$ , we modify the boxing

system around  $\mu^{(1)}$  at  $r$  given in (7.10), so that the  $r$ th box is somewhat smaller and has parameter  $\mu^{(2)}$ . The diameters of boxes and subboxes with index  $r + j, j \geq 0$  in system 1 are equal to those with index  $j \geq 0$  in System-2. Finally we define  $k^*$  in (7.16) as the last index  $k$  that  $\text{Box}_{r+k}^{(1)} \cap \text{Box}_k^{(2)} = \emptyset$  and then we centre  $\text{Box}_{k^*+1}^{(1,2)}$  at  $\lfloor m/2 \cdot \epsilon \rfloor$  to contain both  $\text{Box}_{r+k^*}^{(1)}, \text{Box}_{k^*}^{(2)}$ . Then, the paths  $\pi^{(q)}$  and their connection  $\pi^{\text{conn}}$  described in (7.18), (7.20) exists with probability  $\eta(K)$ , that is at most twice the rhs of (6.17), with  $\mu$  replaced by  $m(K)$ . The length of this connection is at most  $\epsilon_K$ , the rhs of (7.21).

This argument, for a fixed  $K$ , establishes a connection between the best explosive paths  $\pi_{\text{opt}}(0)$  and  $\pi_{\text{opt}}(\lfloor m\epsilon \rfloor)$  that has length at most  $|\pi_{\text{opt}}(0)|_L + \pi_{\text{opt}}(\lfloor m\epsilon \rfloor) + \epsilon_K$ , and is not successful with probability  $\eta(K)$ . To obtain an *almost sure* connection, we repeat this procedure for an (increasing) sequence  $(K_i)_{i \geq 1}$  so that the error probability  $\eta(K_i)$  is summable. Then, by the Borel-Cantelli lemma, it will happen only finitely many times that a connection between  $\tilde{v}^1(K), \tilde{v}^2(K)$  with length at most  $\epsilon_{K_i}$  can *not* be established. Since  $|\pi_{\text{opt}}^q|_L = Y^S(v^q)$ , the length of the total path constructed is at most  $Y^S(0) + Y^S(\lfloor m\epsilon \rfloor) + \epsilon_{K_i}$ , for arbitrarily small  $\epsilon_{K_i} > 0$ . Asymptotic independence of  $Y^S(0), Y^S(\lfloor m\epsilon \rfloor)$  as  $m \rightarrow \infty$  was already part of [51, Theorem 1.7], and can be established in a similar manner as we did for the GIRG, see the proof of (3.21) between (3.25) - (3.29).  $\square$

## 9. HYPERBOLIC RANDOM GRAPHS AND THRESHOLD ASSUMPTION

In this section we discuss the relation between hyperbolic random graphs and GIRGs, and show that Theorem 2.12 is valid for hyperbolic random graphs. For the model description of hyperbolic random graphs, we follow [42] and [63]. Let us denote by  $(\phi_v, r_v)$  the (hyperbolic) angle and radius of a location of a point within a disk of radius  $R$ , and define the hyperbolic distance  $d_H^{(n)}(u, v)$  of two vertices at  $(\phi_u, r_u), (\phi_v, r_v)$  by the equation

$$\cosh(d_H^{(n)}(u, v)) := \cosh(r_u) \cosh(r_v) - \sinh(r_u) \sinh(r_v) \cos(\phi_u - \phi_v). \quad (9.1)$$

**Definition 9.1** (Hyperbolic random graphs). *For parameters  $C_H, \alpha_H, T_H > 0$ , set  $R_n = 2 \log n + C_H$ , and sample  $n$  vertices independently from a circle of radius  $R_n$  so that for each  $v \in [n]$ ,  $\phi_v$  is uniform in  $[0, 2\pi]$ , and  $r_v \in [0, R_n]$  follows a density  $f_n(r) := \alpha_H \sinh(\alpha_H r) / (\cosh(\alpha_H R_n) - 1)$ , independently of  $\phi_v$ . Then two vertices are connected in the threshold hyperbolic random graphs whenever  $d_H^{(n)}(u, v) \leq R_n$ , while in a parametrised version [63, Section VI] they are connected independently of everything else, with probability*

$$p_H^{(n)}(d_H^{(n)}(u, v)) := (1 + \exp\{(d_H^{(n)}(u, v) - R_n)/2T_H\})^{-1}. \quad (9.2)$$

*Let us denote the obtained random graphs by  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  when (9.2) applies and  $\text{HG}_{\alpha_H, C_H}(n)$  when the threshold  $d_H^{(n)}(u, v) \leq R_n$  is applied.*

We show below that the power-law exponent of the degree distribution in HRG is  $\tau_H = 2\alpha_H + 1$ . The most important theorem of this section is the following:

**Theorem 9.2.** *Consider the models  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  and  $\text{HG}_{\alpha_H, C_H}(n)$  with  $\alpha_H \in (1/2, 1)$ . Assign to each present edge  $e$  and edge-length  $L_e$ , and i.i.d. copy of the random variable  $L$ . Then the results in Theorem 2.12 are valid for these models, i.e. the weighted distance between two uniformly chosen vertices in the giant component converges in distribution.*

We show this theorem by deterministically transforming HRGs into a GIRG model, and showing that the assumptions in Theorem 2.12 all hold for the obtained GIRG. The connection to GIRGs is derived as follows: set  $d := 1, \mathcal{X}_1 := [-1/2, 1/2]$ , and let, for a vertex  $v = (\phi_v, r_v)$ ,

$$x_v := (\phi_v - \pi)/(2\pi), \quad W_v^{(n)} := \exp\{(R_n - r_v)/2\}. \quad (9.3)$$

For threshold HRGs, we need to modify Assumptions 2.2, 2.5. So, let us introduce two functions, for not necessarily equal  $\underline{a}_1, \bar{a}_1 > 0, a_2 > 0$  and  $\gamma > 0$ ,

$$\begin{aligned} \bar{g}_\infty(x, w_1, w_2) &:= \mathbb{1}_{\{\|x\| \leq \bar{a}_1(w_1 w_2)^{1/d}\}}, \\ g_\infty(x, w_1, w_2) &:= e^{-a_2((\log w_1)^\gamma + (\log w_2)^\gamma)} \mathbb{1}_{\{\|x\| \leq \underline{a}_1(w_1 w_2)^{1/d}\}}. \end{aligned} \quad (9.4)$$

and consider the following two assumptions (in place of Assumptions 2.2, 2.5):

**Assumption 9.3.** *There exist parameters  $\gamma \in (0, 1)$ ,  $a_1, a_2 \in \mathbb{R}_+$  and  $0 < c_1 \leq C_1 < 1$ , such that for all  $n$ ,  $g_n$  in (2.1) satisfies*

$$c_1 \underline{g}_\infty(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}) \leq g_n(x_u, x_v, (W_i^{(n)})_{i \in [n]}) \leq C_1 \bar{g}_\infty(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}). \quad (9.5)$$

We call a model satisfying Assumption 9.3 a *threshold GIRG*. Assumption 2.5 cannot hold directly for threshold GIRGs, since whenever the limiting function contains an indicator that happens to be 0, the error bound in (2.11) cannot be satisfied. As a result, we modify Assumption 2.5 as follows:

**Assumption 9.4** (Limiting probabilities for threshold GIRG). *Recall Definition 2.1 and set  $(\mathcal{X}_d, \nu) := ([-1/2, 1/2]^d, \mathcal{L}_{eb})$ . We assume that on some event  $\mathcal{L}_n$  that satisfies  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{L}_n) = 1$ , there exists a function  $h : \mathbb{R}^d \times \mathbb{R}_+^2 \rightarrow [0, 1]$  and intervals  $I_\Delta(n) \subseteq \mathbb{R}_+$ ,  $I_w(n) \subseteq [1, \infty)$  such that  $g_n$  in (2.1) satisfies for  $\nu$ -almost every  $x \in \mathcal{X}_d$ , and all fixed  $\Delta, w_u, w_v$  with  $\|\Delta\| \in I_\Delta(n)$ ,  $w_u, w_v \in I_w(n)$ , that for some set  $\mathcal{I}_{w_u, w_v}(n)$  with measure at most  $\epsilon(n)$ ,*

$$|\tilde{g}_n(x, x + \Delta/n^{1/d}, w_u, w_v, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}) - h(\Delta, w_u, w_v)| \leq \mathbb{1}\{\Delta \in (w_u w_v)^{1/d} \mathcal{I}_{w_u, w_v}(n)\}, \quad (9.6)$$

with  $\epsilon(n) \rightarrow 0$ ,  $I_\Delta(n) \rightarrow (0, \infty)$ ,  $I_w(n) \rightarrow [1, \infty)$  as  $n \rightarrow \infty$ .

The next two theorems tell us that weighted distances in threshold GIRGs still converge, and that weighted distances in  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  and  $\text{HG}_{\alpha_H, C_H}(n)$  converge in distribution:

**Proposition 9.5.** *Theorem 2.12 remains valid when Assumptions 2.2 and 2.5 are switched to Assumption 9.3 and 9.4, respectively.*

**Theorem 9.6.** *With the mapping in (9.3),  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  becomes a  $\text{GIRG}_{W, L}(n)$  model that satisfies Assumptions 2.2, 2.4, 2.5 with parameters  $\alpha := 1/T_H$ ,  $\tau := 2\alpha_H + 1$  and  $d = 1$ . Further,  $\text{HG}_{\alpha_H, C_H}(n)$  satisfies Assumptions 9.3, 2.4, 9.4 with parameters  $\tau := 2\alpha_H + 1$  and  $d = 1$ . In particular, Theorem 2.12 is valid for  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  and  $\text{HG}_{\alpha_H, C_H}(n)$ .*

Observe that Theorem 9.2 follows directly from Proposition 9.5 and Theorem 9.6.

*Proof of Theorem 9.6 subject to Proposition 9.5.* We start showing that  $\text{HG}_{\alpha_H, C_H, T_H}(n)$ ,  $\text{HG}_{\alpha_H, C_H}(n)$  satisfy Assumption 2.4. In [23, Lemma 7.2] it is shown that when  $\alpha_H > 1/2$ , the weight distribution follows a (weak) power-law with  $\tau = 2\alpha_H + 1$  in the limit. This is a similar result to showing that the degree distribution follows a power law in the limit, a result in [42]. Assumption 2.4 also requires an investigation on the  $n$ -dependent weight distribution, that we show now. The condition  $r_v \in [0, R_n]$  implies that  $W_v^{(n)} \in [1, e^{C_H/2} n]$ . For any  $1 \leq x \leq e^{C_H/2} n$ ,

$$\begin{aligned} \mathbb{P}(W_v^{(n)} > x) &= \mathbb{P}(r_v < R_n - 2 \log x) = \int_0^{R_n - 2 \log x} f_n(t) dt = \frac{\cosh(\alpha_H(R_n - 2 \log x)) - 1}{\cosh(\alpha_H R_n) - 1} \\ &= \frac{x^{-2\alpha_H} (n^2 e^{C_H})^{\alpha_H} (1 + (x e^{-C_H/2}/n)^{4\alpha_H})/2 - 1}{(n^2 e^{C_H})^{\alpha_H} (1 + (e^{-C_H/2}/n)^{4\alpha_H})/2 - 1}. \end{aligned} \quad (9.7)$$

So, set  $[1, M_n] := [1, e^{C_H/2} n]$  in Assumption 2.4, and then

$$\mathbb{P}(W_v^{(n)} > x) = x^{-2\alpha_H} \ell_n(x), \quad (9.8)$$

where for all  $x \in [1, M_n]$   $1/2 \leq \ell_n(x) \leq 2 + 6e^{-C_H \alpha_H/2}$  holds. Hence, (2.6) is satisfied with  $\underline{\ell} \equiv 1/2$ ,  $\bar{\ell} \equiv 2 + 6e^{-C_H \alpha_H/2}$ . For fixed  $x \in [1, \infty)$ ,  $\mathbb{P}(W_v^{(n)} > x)$  converges to  $\mathbb{P}(W > x) = x^{-2\alpha_H}$ , thus (2.7) is also satisfied with  $\tau := 2\alpha_H + 1$ . Next we show that  $\text{HG}_{\alpha_H, C_H}(n)$  satisfies Assumption 9.3. Let

$$\theta_n(r_u, r_v) := \arg \max_{\phi \leq \pi} \{d_H^{(n)}((0, r_u), (\phi, r_v)) \leq R_n\} = \arccos \left\{ \frac{\cosh(r_u) \cosh(r_v) - \cosh(R_n)}{\sinh(r_u) \sinh(r_v)} \right\} \quad (9.9)$$

be the largest angle of a vertex with radius  $r_v$  that a vertex at  $(0, r_u)$  is connected to. In [42, Lemma 3.1], it is shown that, when  $r_u + r_v \geq R_n$ ,

$$2e^{(R_n - r_u - r_v)/2} (1 + \Theta(e^{R_n - r_u - r_v})) \leq \theta_n(r_u, r_v) \leq 2e^{(R_n - r_u - r_v)/2} (1 + \Theta(e^{-2r_u} + e^{-2r_v})). \quad (9.10)$$



On the other hand, when  $r_u + r_v \leq R_n$ , then  $\theta_n(r_u, r_v) = \pi$  due to the triangle inequality. Note that  $r_u + r_v \leq R_n$  is equivalent to  $w_u w_v / n \geq e^{C_H/2}$ . Using (9.3), this corresponds to that for some  $c_1, C_1, \underline{a}_1, \bar{a}_1 > 0$ ,

$$\begin{aligned} \|x_u - x_v\| &\leq c_1 \min\{1, \underline{a}_1 w_u w_v / n\} \Rightarrow d_H(u, v) \leq R_n, \\ \|x_u - x_v\| &\geq C_1 \min\{1, \bar{a}_1 w_u w_v / n\} \Rightarrow d_H(u, v) \geq R_n. \end{aligned} \quad (9.11)$$

which is also shown in [23, Lemma 7.5]. Hence, Assumption 9.3 is satisfied. Next we show that Assumption 2.2 is satisfied for  $\text{HG}_{\alpha_H, C_H, T_H}(n)$ . Let  $s_n := d_H^{(n)}((x_u, w_u), (x_v, w_v)) - R_n$ . Note that

$$e^{s_n} = 2 \cosh(d_H^{(n)}) e^{-R_n} + e^{-s_n} e^{-R_n} = 2 \cosh(d_H^{(n)}) e^{-R_n} + e^{-s_n} \Theta(n^{-4}). \quad (9.12)$$

By noting that  $r_q = R_n - 2 \log w_q$ ,  $\phi_q = 2\pi x_q + \pi$  for  $q \in \{u, v\}$ , using a trigonometric identity  $\cosh(x - y) = \cosh(x) \cosh(y) - \sinh(x) \sinh(y)$ ,

$$\begin{aligned} \cosh(d_H^{(n)}) e^{-R_n} &= \cosh(r_u - r_v) e^{-R_n} + (1 - \cos(2\pi(x_u - x_v))) \sinh(r_u) \sinh(r_v) e^{-R_n} \\ &= \frac{(w_u/w_v)^2 + (w_v/w_u)^2}{n^2} e^{-C_H} \\ &\quad + 2\pi^2 \|x_u - x_v\|^2 \left(1 + \Theta(\|x_u - x_v\|^2)\right) \frac{e^{C_H} n^2}{4w_u^2 w_v^2} \left(1 - \frac{e^{-2C_H} w_u^4}{n^4}\right) \left(1 - \frac{e^{-2C_H} w_v^4}{n^4}\right). \end{aligned} \quad (9.13)$$

Since  $p_H^{(n)}(d_H^{(n)}) = (1 + e^{s_n/(2T_H)})^{-1}$ , from a combination of (9.12) and (9.13), using that  $(1 \wedge x^{-1})/2 \leq (1 + x)^{-1} \leq 1 \wedge x^{-1}$ , it follows that for any angle difference  $\|x_u - x_v\|$ ,

$$c_1 \left(1 \wedge \underline{a}_1 \left(\frac{w_u w_v}{n \|x_u - x_v\|}\right)^{1/T_H}\right) \leq p_H^{(n)}(d_H^{(n)}) \leq C_1 \left(1 \wedge \bar{a}_1 \left(\frac{w_u w_v}{n \|x_u - x_v\|}\right)^{1/T_H}\right), \quad (9.14)$$

that is,  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  satisfies Assumption 2.2 with  $\tau = 2\alpha_H + 1$ ,  $\alpha = 1/T_H$  as well. This has appeared in [23, Lemma 7.5]. It remains to show that Assumption 2.5 is satisfied for  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  and Assumption 9.4 is satisfied for  $\text{HG}_{\alpha_H, C_H}(n)$ . Note that the original GIRG model allows for the connection probabilities to depend on the whole collection of weights  $(W_i)_{i \in [n]}$ , while in  $\text{HG}_{\alpha_H, C_H, T_H}(n)$  and  $\text{HG}_{\alpha_H, C_H}(n)$  this is actually not the case, connection probabilities only depend on  $w_u, w_v$  and  $x_u - x_v$  per definition. We need to show that a limiting connection probability  $h(\Delta, w_u, w_v)$  between vertices with  $u = (x_u, w_u)$ ,  $v = (x_u + \Delta/n, w_v)$  exists when we set  $w_u, w_v$  to be fixed *constants*, and that (2.11) holds.

We continue showing that the limiting function  $h(\Delta, w_u, w_v)$  exists when we set  $x_u - x_v = \Delta/n$ . Indeed, combining (9.12) with twice the rhs of (9.13) equals then, (since  $w_u, w_v \geq 1$ )

$$e^{s_n} = \frac{\Delta^2 e^{C_H} \pi^2}{w_u^2 w_v^2} \left(1 + \Theta(\Delta^2/n^2) + \Theta((w_u^4 + w_v^4)/n^4)\right) + \Theta\left(\frac{w_u^2 + w_v^2}{n^2}\right) + \Theta\left(\frac{w_u^2 w_v^2}{\Delta^2 n^4}\right). \quad (9.15)$$

Thus, its limit exists for every fixed  $\Delta, w_u, w_v$  with  $e^s = \Delta^2 e^{C_H} \pi^2 / (w_u^2 w_v^2)$ . Hence,

$$p_H^{(n)}(d_H^{(n)}) \rightarrow (1 + e^{s/(2T_H)})^{-1} = (1 + (e^{C_H/2} \|\Delta\| \pi / (w_u w_v))^{1/T_H})^{-1} =: h(\Delta, w_u, w_v), \quad (9.16)$$

while for the threshold case,

$$\mathbb{1}\{d_H^{(n)} \leq R_n\} \rightarrow \mathbb{1}\{s(\Delta, w_u, w_v) \leq 0\} = \mathbb{1}\{\|\Delta\| \leq e^{-C_H/2} w_u w_v / \pi\} =: h_\infty(\Delta, w_u, w_v). \quad (9.17)$$

From (9.16) and (9.17), it follows that the limiting functions  $h, h_\infty$  also satisfy Assumptions 2.2 and 9.3, respectively. What is left is to show that (2.11) in Assumption 2.5 is satisfied. For this we need to estimate the relative difference between the connection probabilities and their limit. Let  $p(x) := (1 + x^{1/(2T_H)})^{-1}$ . Then  $p_H^{(n)}(d_H^{(n)}) = p(e^{s_n})$ ,  $h(\Delta, w_u, w_v) = p(e^s)$ . So,

$$\begin{aligned} |p_H^{(n)}(d_H^{(n)}) - p(e^s)| / p(e^s) &= |p'(e^s)(e^{s_n} - e^s)| / p(e^s) + o(e^{s_n} - e^s) / p(e^s) \\ &= p(e^s) e^{s/(2T_H)} |e^{s_n} / e^s - 1| + o((e^{s_n} - e^s) / p(e^s)). \end{aligned} \quad (9.18)$$

since  $p'(e^s) = -e^{s(1/(2T_H)-1)} p(e^s)^2 / (2T_H)$ . Using (9.15), and that  $e^s = \Theta(\Delta^2 / (w_u^2 w_v^2))$ ,

$$|e^{s_n} / e^s - 1| \leq \Theta(\Delta^2 / n^2 + (w_u^4 + w_v^4) / n^4 + w_u^2 w_v^2 (w_u^2 + w_v^2) / (\Delta^2 n^2) + w_u^4 w_v^4 / \Delta^4 n^4). \quad (9.19)$$

Finally, returning to (9.18), and recalling that  $1/T_H =: \alpha$ ,  $e^{s/(2T_H)} = \Theta(\Delta/w_u w_v)^\alpha$  and so  $p(e^s) \leq \min\{1, \Theta(w_u w_v/\Delta)^\alpha\}$ ,

$$\begin{aligned} |p_H^{(n)}(d_H^{(n)}) - p(e^s)|/p(e^s) &\leq \min\{1, \Theta(w_u w_v/\Delta)^\alpha\} \Theta((\Delta/w_u w_v)^\alpha) \\ &\quad \cdot \Theta(\Delta^2/n^2 + (w_u^4 + w_v^4)/n^4 + w_u^2 w_v^2 (w_u^2 + w_v^2)/(\Delta^2 n^2) + w_u^4 w_v^4/(\Delta^4 n^4)) \\ &= \Theta(\Delta^2/n^2 + (w_u w_v/\Delta)^{4-\alpha}/n^4 + (w_u w_v/\Delta)^{2-\alpha} (w_u^2 + w_v^2)/n^2) \end{aligned} \quad (9.20)$$

Note that the error terms are at most  $\Theta(n^{-1})$  whenever  $\|\Delta\| \in [n^{-\beta(\alpha)}, n^{\beta(\alpha)}]$ ,  $w_u, w_v \in [1, n^{\beta(\alpha)}]$ , where

$$\beta(\alpha) = \mathbb{1}\{\alpha < 2\} \frac{1}{3(2-\alpha) + 2} + \mathbb{1}\{\alpha \geq 2\} \frac{1}{\alpha}. \quad (9.21)$$

Thus, Assumption 2.5 is satisfied with  $\epsilon(n) = n^{-1}$  and  $I_\Delta(n) = [n^{-\beta(\alpha)}, n^{\beta(\alpha)}]$ ,  $I_w(n) := [1, n^{\beta(\alpha)}]$ . We continue showing that Assumption 9.4 holds for  $\text{HG}_{\alpha_H, C_H}(n)$ . Combining (9.15) and (9.17),

$$\mathbb{1}\{d_H^{(n)} \leq R_n\} = \mathbb{1}\{e^{s_n} \leq 1\} = \mathbb{1}\left\{\|\Delta\| \leq \frac{w_u w_v}{e^{C_H/2\pi}} \cdot \frac{1 + \Theta(w_u^2 w_v^2/(\Delta^2 n^4) + (w_u^2 + w_v^2)/n^2)}{1 + \Theta(\Delta^2/n^2) + \Theta((w_u^4 + w_v^4)/n^4)}\right\}, \quad (9.22)$$

and hence

$$\mathbb{1}\{d_H^{(n)} \leq R_n\} - \mathbb{1}\{\|\Delta\| \leq \frac{w_u w_v}{e^{C_H/2\pi}}\} \leq \mathbb{1}\{\|\Delta\| \in w_u w_v \mathcal{I}_{w_u, w_v}(n)\}, \quad (9.23)$$

where  $\mathcal{I}_{w_u, w_v}(n)$  is an interval of length

$$|\mathcal{I}_{w_u, w_v}(n)| = \frac{1}{e^{C_H/2\pi}} \Theta(w_u^2 w_v^2/(\|\Delta\|^2 n^4) + (w_u^2 + w_v^2)/n^2 + \|\Delta\|^2/n^2 + (w_u^4 + w_v^4)/n^4). \quad (9.24)$$

Note that the length of this interval is at most  $1/n$  whenever  $\|\Delta\| \in [n^{-1/2}, n^{1/2}]$ ,  $w_u, w_v \in [1, n^{1/2}]$ . Thus, Assumption 9.4 is satisfied with  $\epsilon(n) := 1/n$ ,  $I_\Delta(n) := [n^{-1/2}, n^{1/2}]$ ,  $I_w(n) := [1, n^{1/2}]$ . We comment that for HRGs the underlying space is the one-dimensional *torus*  $[-1/2, 1/2)$  instead of  $[-1/2, 1/2]$ . This is not an issue since in the proofs we made sure the paths constructed does not touch the boundary of  $\tilde{\mathcal{A}}_d(n)$ , that is equivalent to not wrapping around the torus.  $\square$

The content of the following claim is the modified version of Claim 3.4 for threshold GIRGs.

**Claim 9.7.** *Suppose that Assumption 9.4 holds instead of Assumption 2.5. Then, Claim 3.4 remains valid with a different proof.*

*Proof.* By the coupling described in Claim 3.3, the vertex set of  $\text{BGIRG}_{W,L}(n)$  is a subset of  $\mathcal{V}_{1+\xi_n}$ . We describe the new coupling of the edges. Following the proof of Claim 3.4, the vertex-weights are not equal in  $\text{BGIRG}_{W,L}(n)$  and in  $\text{EGIRG}_{W,L}(1+\xi_n)$  with probability at most  $\epsilon_{\text{TV}}(n)$ , i.e.

$$\mathbb{P}(W_{i_1}^{(n)} \neq W_{i_1} \text{ or } W_{i_2}^{(n)} \neq W_{i_2}) = 2\epsilon_{\text{TV}}(n). \quad (9.25)$$

Given that the two pairs of weights are equal and equal to  $W_{i_1}, W_{i_2}$ , the edge is present in  $\bar{\mathcal{E}}_\lambda$  when  $\|x_{i_1} - x_{i_2}\| \leq \bar{a}_1(W_{i_1} W_{i_2})^{1/d}$ , and, by (9.5) that bounds the presence of the same edge in  $\mathcal{E}_B(n)$  (note that the  $n$ -dependence in (9.5) disappears when moving from  $\text{GIRG}_{W,L}(n)$  to  $\text{BGIRG}_{W,L}(n)$ ). Hence,  $\mathcal{E}_B(n) \subseteq \bar{\mathcal{E}}_{1+\xi_n}$  holds. We continue coupling  $\mathcal{E}_B(n)$  and  $\mathcal{E}_{1+\xi_n}$ . A de-coupling happens precisely when for a possible edge  $e$  connecting vertices with locations and weights  $(x_{i_1}, W_{i_1}), (x_{i_2}, W_{i_2})$ ,  $\|x_{i_1} - x_{i_2}\| \in (W_{i_1} W_{i_2})^{1/d} \mathcal{I}_{W_{i_1} W_{i_2}}(n)$ . Thus, conditioned on the vertex weights and on the location  $x_{i_1}$ ,

$$\begin{aligned} \mathbb{P}(x_{i_2} - x_{i_1} \in (W_{i_1} W_{i_2})^{1/d} \mathcal{I}_{W_{i_1} W_{i_2}}(n) \mid \|x_{i_2} - x_{i_1}\| \leq \bar{a}_1(W_{i_1} W_{i_2})^{1/d}, E_{\Delta, W}, W_{i_1}, W_{i_2}, x_{i_1}) \\ \leq (W_{i_1} W_{i_2} \epsilon(n)) / (\bar{a}_1^d W_{i_1} W_{i_2}) \leq \epsilon(n) / \bar{a}_1^d. \end{aligned} \quad (9.26)$$

The condition that the event  $E_{\Delta, W}$  holds (see Claim 3.4) guarantees that this bound is true. As a result, (3.9) remains valid with  $2\epsilon_{\text{TV}}(n) + \epsilon(n)/\bar{a}_1^d$  instead of  $\epsilon(n)$  and a union bound finishes the proof.  $\square$

*Proof of Theorem 9.5.* We discuss how the lemmas, claims and propositions in this paper need to be adjusted when Assumptions 9.3, 9.4 hold instead of Assumptions 2.2, 2.5. Let us equip the  $\text{EGIRG}_{W,L}(\lambda)$  model with edge probability  $h_\infty$ , such that  $\tilde{g}_n$  satisfies Assumption 2.5 with  $h := h_\infty$ . This implies that Claim 2.6 holds for  $h_\infty$ , i.e. there exist constants  $\underline{c}_1, \overline{C}_1$ , with  $0 < \underline{c}_1 \leq c_1 \leq C_1 \leq \overline{C}_1 < \infty$ , such that

$$\underline{c}_1 \underline{g}_\infty(\Delta, w_u, w_v) \leq h_\infty(\Delta, w_u, w_v) \leq \overline{C}_1 \overline{g}_\infty(\Delta, w_u, w_v). \quad (9.27)$$

Then, equip the  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  model with the edge probability  $\overline{C}_1 \overline{g}_\infty$ . We have seen in Claim 9.7 that the coupling of edges in (3.4) remains valid. When changing the definition of the Bernoulli BRW in (4.1) to  $N_x^B(\underline{i}) \stackrel{d}{=} \overline{C}_1 \overline{g}_\infty(x_{\underline{i}} - x, W_{\underline{i}}, W_x)$ , Lemma 4.1 and its extension in (4.6) hold, since in the coupling it is only used that (4.3) holds between  $\text{BerBRW}_\lambda$  and  $\overline{\text{EGIRG}}_{W,L}(\lambda)$ , and the connection probability is at least as much in  $\overline{\text{EGIRG}}_{W,L}(\lambda)$  as in  $\text{EGIRG}_{W,L}(\lambda)$  between two vertices with given weights and locations, which stays true under Assumption 9.3. Similarly, in the proof of Theorem 2.11, if one replaces the rv  $\text{Ber}(\overline{C}_1 \min\{1, \|y\|^{-\alpha d}(ws)^\alpha\})$  in (4.12) by  $\mathbb{1}_{\{\|y\| \leq \overline{a}_1(ws)^{1/d}\}}$ , and in (4.13) drops the second sum and changes the constraint in the first sum to  $y \in \mathcal{V}_\lambda : \|y\| \leq \overline{a}_1(ws)^{1/d}$ , then

$$\mu(w, ds) \leq F_W(ds) \lambda \overline{a}_1^d \text{Vol}_d ws, \quad (9.28)$$

where  $\text{Vol}_d$  is the volume of the d-dimensional unit ball. When we adjust  $c_\lambda$  in (4.15) to  $\lambda \overline{a}_1^d \text{Vol}_d$ , the rest of the proof follows analogously. Hence, Theorem 2.11 and thus Corollary 4.2 follow. Likewise, Proposition 4.4 holds, as Claim 4.6 is satisfied. In the proof of Claim 4.6, (4.23) can be reduced to the second integral, changing the constraint in the sum to  $\|x\| \leq \overline{a}_1(ws)^{1/d}$ , which yields

$$\mathbb{E}[N_w(\geq 1, \geq S)] \leq \overline{a}_1^d \lambda \text{Vol}_d C_\tau w^{\tau-1} S^{-d(\tau-2)} \ell(S^d/w), \quad (9.29)$$

and similarly,

$$\mathbb{E}[N_w(\geq u, \geq 0)] \leq \overline{a}_1^d \lambda \text{Vol}_d C_\tau w u^{-(\tau-2)} \ell(u). \quad (9.30)$$

These yield that Proposition 3.6 holds. For the weight-dependent percolation, as defined in Definition 6.1, we show that Claim 6.2 holds, with  $(W_{p,i}^{(n)})_{i \in [n]} := (W_i^{(n)})_{i \in [n]}$ , i.e. unchanged weights. By definition, an edge  $e = (u, v)$  is present with probability

$$\begin{aligned} \mathbb{P}((u, v) \in \mathcal{E}(G^{\text{thr}}) \mid (x_i, W_i^{(n)})_{i \in [n]}) &= p(W_u^{(n)}, W_v^{(n)}) \tilde{g}_n^B(x_u, x_v, W_u^{(n)}, W_v^{(n)}, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}) \\ &=: \tilde{g}_n^{\text{thr}}(x_u, x_v, W_u^{(n)}, W_v^{(n)}, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}), \end{aligned} \quad (9.31)$$

which we can bound from above and below using Assumption 9.3, by

$$\begin{aligned} \tilde{g}_n^{\text{thr}}(x_u, x_v, W_u^{(n)}, W_v^{(n)}, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}) &\leq C_1 \overline{g}_\infty(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}), \\ \tilde{g}_n^{\text{thr}}(x_u, x_v, W_u^{(n)}, W_v^{(n)}, (W_i^{(n)})_{i \in [n] \setminus \{u, v\}}) &\geq c_1 p(W_u^{(n)}, W_v^{(n)}) \underline{g}_\infty(n^{1/d}(x_u - x_v), W_u^{(n)}, W_v^{(n)}). \end{aligned} \quad (9.32)$$

When  $p(w_u, w_v)$  satisfies (6.2),

$$p(w_u, w_v) \exp\{-a_2(\log w_u)^\gamma - a_2(\log w_v)^\gamma\} \geq \exp\{-(a_2 + c)(\log w_u)^{\gamma_{\max}} - (a_2 + c)(\log w_v)^{\gamma_{\max}}\},$$

with  $\gamma_{\max} = \max\{\gamma, \tilde{\gamma}\}$ . So, (9.4) is satisfied again with new constant  $a_2 + c$  in place of  $a_2$ , and new exponent  $\gamma_{\max} \in (0, 1)$  in place of the original  $\gamma$ . Hence we can define  $\underline{G}^p \subseteq G^{\text{thr}}$  as a new threshold GIRG with the edge probabilities

$$\begin{aligned} g_n^p(x_u, x_v, W_u^{(n)}, W_v^{(n)}) &:= \exp\{-(a_2 + c)(\log W_u^{(n)})^{\gamma_{\max}} - (a_2 + c)(\log W_v^{(n)})^{\gamma_{\max}}\} \\ &\quad \cdot c_1 \mathbb{1}_{\{\|x\| \leq \underline{a}_1(W_u^{(n)} W_v^{(n)})^{1/d}\}}, \end{aligned} \quad (9.33)$$

proving together with (9.32) that the percolated graph contains a subgraph that is again an instance of the threshold GIRG, with the same parameters and the same weights  $(W_i^{(n)})_{i \in [n]}$ . So, Claim 6.2 is valid. Turning to Lemma 6.3, the proof of the events  $F_k^{(1)}$  only concerns weights, so it holds unchanged. Investigating the events  $F_k^{(2)}$ , using (6.12) and the events  $F_k^{(1)}$ , the lower bound on the edge probabilities in (6.27) is satisfied for this model as well, as

$$\|c_k^{(i)} - c_{k+1}^{(j)}\|^d / (W_{c_k^{(i)}}^{(n)} W_{c_{k+1}^{(j)}}^{(n)}) \leq d^{d/2} \mu^{C^k(CD - (1-\delta)(1+C)/(\tau-1))}, \quad (9.34)$$

which decreases with  $\mu$ . From here the proof could be followed word-by-word. Thus, Lemma 6.3 holds as well. It directly follows that Proposition 3.5 is satisfied as well. Finally, since [22, Lemma 5.5] holds for the threshold GIRG model as well, as shown in said paper, the proofs of Claims 7.1, 7.2 and 7.3 follow for threshold GIRGs. Hence, Lemma 3.7 and Proposition 3.8 are satisfied, finishing the proof.  $\square$

**Acknowledgements.** The work of JK is supported by the Netherlands Organisation for Scientific Research (NWO) through VENI grant 639.031.447 (JK). We thank the unknown reviewers for their comments that helped to improve the presentation of the paper.

## REFERENCES

- [1] M. A. Abdullah, M. Bode, and N. Fountoulakis. Typical distances in a geometric model for complex networks. *Internet Mathematics 1 (electronic)*, August 2017.
- [2] E. Adriaans and J. Komjáthy. Weighted distances in scale-free configuration models. *Journal of Statistical Physics*, 173(3):1082–1109, Nov 2018.
- [3] W. Aiello, A. Bonato, C. Cooper, J. Janssen, and P. Prałat. A spatial web graph model with local influence regions. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4863, pages 96–107, 2007.
- [4] R. Albert and A. Barabási. Statistical mechanics of complex networks. *Reviews Of Modern Physics*, 74(1):47–97, January 2002.
- [5] R. Albert, H. Jeong, and A. Barabási. Internet: Diameter of the World-Wide Web. *Nature*, 401(6749), September 1999.
- [6] O. Amini, L. Devroye, S. Griffiths, and N. Olver. On explosions in heavy-tailed branching random walks. *Annals of Probability*, 41(3 B):1864–1899, May 2013.
- [7] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna. Four degrees of separation. In *Proceedings of the 4th Annual ACM Web Science Conference*, pages 33–42. ACM, 2012.
- [8] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science (New York, N.Y.)*, 286(5439), October 1999.
- [9] E. Baroni, R. v. d. Hofstad, and J. Komjáthy. Non-universality of weighted random graphs with infinite variance degree. *Journal of Applied Probability*, 54(1):146–164, 2017.
- [10] E. Baroni, R. v. d. Hofstad, and J. Komjáthy. Tight fluctuations of weight-distances in random graphs with infinite-variance degrees. *Journal of Statistical Physics*, Jan 2019.
- [11] R. Bellman and T. Harris. On age-dependent binary branching processes. *Annals of Mathematics*, 55(2):280–295, March 1952.
- [12] S. Bhamidi, R. v. d. Hofstad, and G. Hooghiemstra. Universality for first passage percolation on sparse random graphs. *Ann. Probab.*, 45(4):2568–2630, July 2017.
- [13] S. Bhamidi, R. Van Der Hofstad, and G. Hooghiemstra. First passage percolation on random graphs with finite mean degrees. *The Annals of Applied Probability*, 20(5):1907–1965, October 2010.
- [14] S. Bhamidi, R. Van der Hofstad, and G. Hooghiemstra. First passage percolation on the Erdős–Rényi random graph. *Combinatorics, Probability and Computing*, 20(5):683–707, 2011.
- [15] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.
- [16] M. Biskup. On the scaling of the chemical distance in long-range percolation models. *Ann. Probab.*, 32(4):2938–2977, 10 2004.
- [17] M. Biskup and J. Lin. Sharp asymptotic for the chemical distance in long-range percolation. arXiv:1705.10380 [math.PR], 2017.
- [18] M. Bode, N. Fountoulakis, and T. Müller. On the largest component of a hyperbolic model of complex networks. *The Electronic Journal of Combinatorics*, 22:1–43, 2015.
- [19] M. Boguñá, F. Papadopoulos, and D. Krioukov. Sustaining the internet with hyperbolic mapping. *Nature Communications*, 1(6), September 2010.
- [20] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, December 1980.
- [21] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, August 2007.
- [22] K. Bringmann, R. Keusch, and J. Lengler. Average distance in a general class of scale-free networks with underlying geometry. *arXiv preprint arXiv:1602.05712*, February 2016.
- [23] K. Bringmann, R. Keusch, and J. Lengler. Sampling Geometric Inhomogeneous Random Graphs in Linear Time. In K. Pruhs and C. Sohler, editors, *25th Annual European Symposium on Algorithms (ESA 2017)*, volume 87 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 20:1–20:15, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

- [24] T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6):1377–1397, September 2006.
- [25] E. Candellero and N. Fountoulakis. Bootstrap percolation and the geometry of complex networks. *Stochastic Processes and their Applications*, 126(1):234 – 264, 2016.
- [26] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences of the United States of America*, 99(25), December 2002.
- [27] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, November 2002.
- [28] Columbia University Medical Center. <http://newsroom.cumc.columbia.edu/blog/2016/07/18/toxic-alzheimers-protein-spreads-brain-via-extracellular-space/>, 2017. [online; accessed 25 august 2017].
- [29] C. Cooper, A. Frieze, and P. Pralat. Some typical properties of the spatial preferred attachment model. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7323, pages 29–40, 2012.
- [30] E. Coupechoux and M. Lelarge. How clustering affects epidemics in random networks. *Advances in Applied Probability*, 46(4):985 – 1008, 2014.
- [31] P. L. Davies. The simple branching process: a note on convergence when the mean is infinite. *J. Appl. Probab.*, 15(3):466–480, 1978.
- [32] M. Deijfen, R. v. d. Hofstad, and G. Hooghiemstra. Scale-free percolation. *Annales de l’institut Henri Poincaré (B) Probability and Statistics*, 49(3):817–838, August 2013.
- [33] P. Deprez, R. S. Hazra, and M. V. Wüthrich. Inhomogeneous long-range percolation for real-life network modeling. *Risks*, 3(1):1–23, January 2015.
- [34] P. Deprez and M. V. Wüthrich. Scale-free percolation in continuum space. *arXiv:1312.1948*, 2013.
- [35] S. Dommers, R. Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. *Journal of Statistical Physics*, 139(1):72–107, April 2010.
- [36] S. N. Dorogovtsev and J. F. F. Mendes. Evolution of networks. *Advances in Physics*, 51(4):1079–1187, June 2002.
- [37] M. Eckhoff, J. Goodman, R. v. d. Hofstad, and F. Nardi. Short paths for first passage percolation on the complete graph. *Journal of Statistical Physics*, 151(6):1056–1088, June 2013.
- [38] M. Eckhoff, J. Goodman, R. van der Hofstad, and F. R. Nardi. Long paths in first passage percolation on the complete graph I. Local PWIT dynamics. *arXiv preprint arXiv:1512.06152*, December 2015.
- [39] M. Eckhoff, J. Goodman, R. van der Hofstad, and F. R. Nardi. Long paths in first passage percolation on the complete graph II. Global branching dynamics. *arXiv preprint arXiv:1512.06145*, December 2015.
- [40] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *The Structure and Dynamics of Networks*, pages 195–206. Princeton University Press, October 2011.
- [41] N. Fountoulakis and T. Müller. Law of large numbers for the largest component in a hyperbolic model of complex networks. *Ann. Appl. Probab.*, 28(1):607–650, 02 2018.
- [42] L. Gugelmann, K. Panagiotou, and U. Peter. Random hyperbolic graphs: degree sequence and clustering. In *International Colloquium on Automata, Languages, and Programming*, pages 573–585. Springer, 2012.
- [43] J. M. Hammersley and D. J. A. Welsh. First-passage percolation, subadditive processes, stochastic networks, and generalized renewal theory. In *Proc. Internat. Res. Semin., Statist. Lab., Univ. California, Berkeley, Calif*, pages 61–110. Springer-Verlag, New York, 1965.
- [44] T. E. Harris. *The theory of branching processes*. Dover Phoenix Editions. Dover Publications, Inc., Mineola, NY, 2002. Corrected reprint of the 1963 original [Springer, Berlin; MR0163361 (29 #664)].
- [45] M. Heydenreich, T. Hulshof, and J. Jorritsma. Structures in supercritical scale-free percolation. *Ann. Appl. Probab.*, 27(4):2569–2604, August 2017.
- [46] C. Hirsch. From heavy-tailed boolean models to scale-free gilbert graphs. *Braz. J. Probab. Stat.*, 31(1):111–143, 02 2017.
- [47] R. v. d. Hofstad. *Random graphs and complex networks. Vol. 1*. Cambridge Series in Statistical and Probabilistic Mathematics, [43]. Cambridge University Press, Cambridge, 2017.
- [48] R. v. d. Hofstad, G. Hooghiemstra, and P. v. Mieghem. First-passage percolation on the random graph. *Probability in the Engineering and Informational Sciences*, 15(2):225–237, 2001.
- [49] R. v. d. Hofstad, G. Hooghiemstra, and P. v. Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, August 2005.
- [50] R. v. d. Hofstad, G. Hooghiemstra, and D. Znamenski. Distances in random graphs with finite mean and infinite variance degrees. *Electronic Journal Of Probability*, 12:703–766, May 2007.
- [51] R. v. d. Hofstad and J. Komjáthy. Explosion and distances in scale-free percolation. *arXiv preprint arXiv:1706.02597*, June 2017.
- [52] R. v. d. Hofstad and J. S. H. v. Leeuwaarden. Hierarchical configuration model. *Internet Mathematics*, January 2017.
- [53] C. D. Howard. Models of first-passage percolation. In *Probability on discrete structures*, volume 110 of *Encyclopaedia Math. Sci.*, pages 125–173. Springer, Berlin, 2004.
- [54] E. Jacob and P. Mörters. Spatial preferential attachment networks: power laws and clustering coefficients. *Annals of Applied Probability*, 25(2):632–663, April 2015.

- [55] P. Jagers. General branching processes as Markov fields. *Stochastic Processes and their Applications*, 32(2):183–212, 1989.
- [56] S. Janson. One, two and three times  $\log n/n$  for paths in a complete graph with random weights. *Combinatorics, Probability and Computing*, 8(4):347–361, 1999.
- [57] S. Janson. Asymptotic equivalence and contiguity of some random graphs. *Random Structures & Algorithms*, 36(1):26–45, January 2010.
- [58] S. Janson. Asymptotic equivalence and contiguity of some random graphs. *Random Structures Algorithms*, 36(1):26–45, 2010.
- [59] P. Kaluza, A. Kölsch, M. T. Gastner, and B. Blasius. The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7(48):1093–1103, July 2010.
- [60] C. Koch and J. Lengler. Bootstrap percolation on geometric inhomogeneous random graphs. *arXiv:1603.02057*, 2016.
- [61] I. Kolossvary and J. Komjathy. First passage percolation on inhomogeneous random graphs. *Advances in Applied Probability*, 47(2), June 2015.
- [62] J. Komjathy. Explosive Crump-Mode-Jagers branching processes. *arXiv:1602.01657 [math.PR]*, 2016.
- [63] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguna. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82:036106, Sep 2010.
- [64] N. Mantel. The detection of disease clustering and a generalized regression approach. *Cancer Research*, 27:209–220, 1967.
- [65] R. Meester and R. Roy. *Continuum percolation*, volume 119. Cambridge University Press, 1996.
- [66] J. M. Montoya and R. V. Sole. Small world patterns in food webs. *Journal of Theoretical Biology*, 214(3):405–412, February 2002.
- [67] T. Muller and M. Staps. The diameter of KPKVB random graphs. *Advances in Applied Probability*, 2019+.
- [68] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.
- [69] M. E. J. Newman. Random graphs with clustering. *Phys. Rev. Lett.*, 103:058701, Jul 2009.
- [70] M. E. J. Newman and J. Park. Why social networks are different from other types of networks. *Physical Review E*, 68(3), September 2003.
- [71] I. Norros and H. Reittu. On a conditionally Poissonian graph process. *Advances in Applied Probability*, 38(1):59–75, March 2006.
- [72] M. . Serrano and M. Boguna. Clustering in complex networks. I. General formalism. *Physical Review E*, 74(5):056114, 2006.
- [73] R. T. Smythe and J. C. Wierman. *First-passage percolation on the square lattice*, volume 671. Springer, 2006.
- [74] J. Travers and S. Milgram. The small world problem. *Psychology Today*, 1(1):61–67, 1967.
- [75] D. J. Watts. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press, 1999.
- [76] D. J. Watts and S. H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), June 1998.
- [77] J. E. Yukich. Ultra-small scale-free geometric networks. *Journal of Applied Probability*, 43(3):665–678, September 2006.

JK: DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, EINDHOVEN UNIVERSITY OF TECHNOLOGY, P.O. BOX 513, 5600 MB EINDHOVEN, THE NETHERLANDS.

BL: DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF BATH, CLAVERTON DOWN, BATH, BA2 7AY, UNITED KINGDOM.

*Email address:* j.komjathy@tue.nl, b.lodewijks@bath.ac.uk